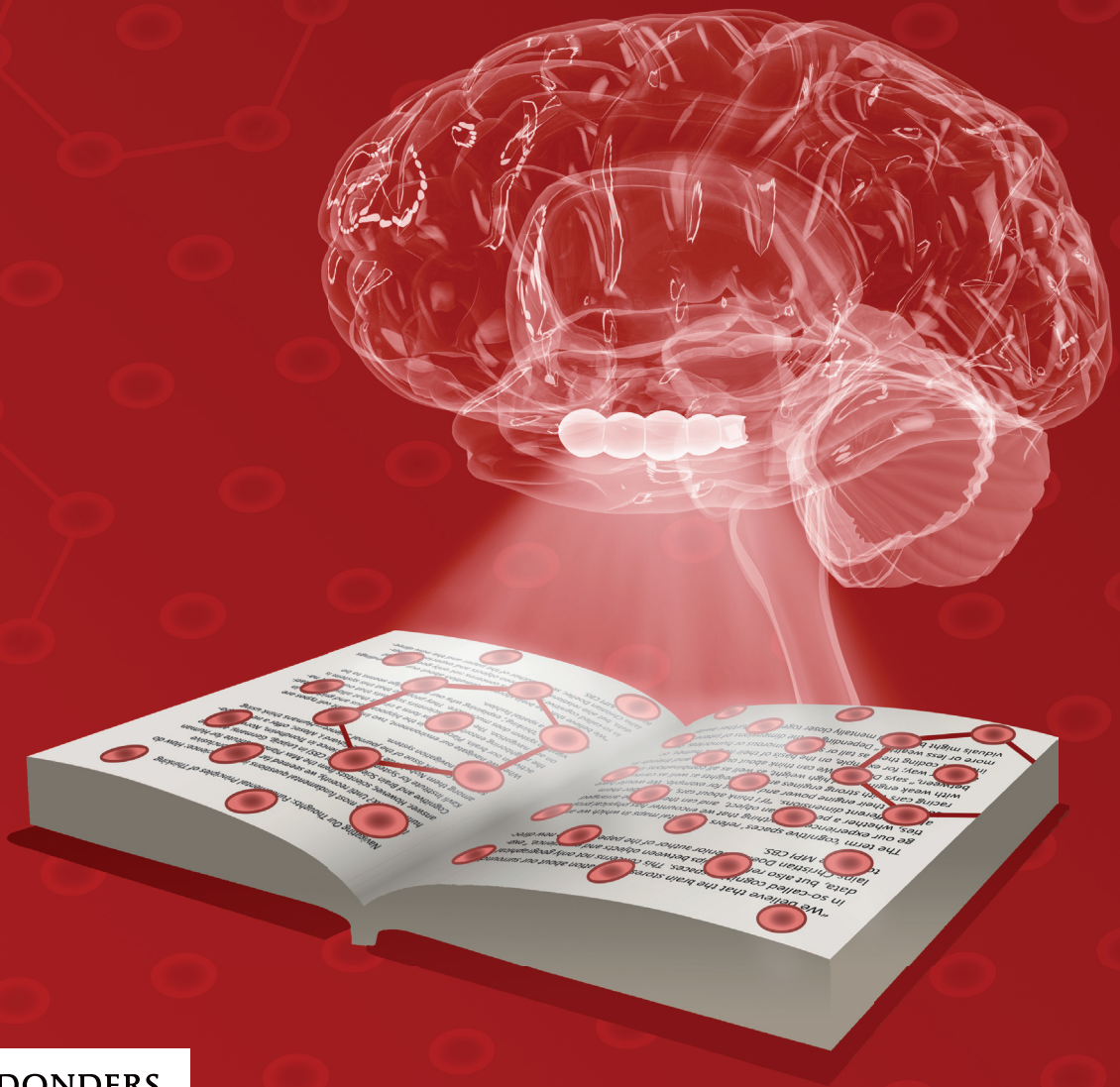
HIPPOCAMPAL-ENTORHINAL CODES FOR SPACE, TIME AND COGNITION

Jacob L. S. Bellmund



DONDERS SERIES

Hippocampal-entorhinal codes for space, time and cognition

Jacob L. S. Bellmund

The work described in this thesis was carried out at the Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen.

ISBN 978 94 6284 271 7

Cover Design MPI CBS and Ella Maru Studio
Layout and Print Ridderprint | www.ridderprint.nl

Hippocampal-entorhinal codes for space, time and cognition

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op
woensdag 9 oktober 2019
om 12:30 uur precies

door

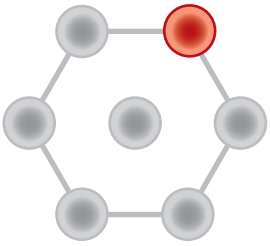
Jacob Lukas Sarid Bellmund
geboren op 1 januari 1988
te Adelebsen (Duitsland)

Promotor	Prof. dr. Christian Döller *Norges teknisk-naturvitenskapelige universitet, Trondheim, Noorwegen*
Copromotor	Dr. Lorena Deuker *Ruhr-Universität Bochum, Duitsland*
Manuscriptcommissie	Prof. dr. Guillén Fernández Dr. Freyja Ólafsdóttir Prof. dr. Hugo Spiers *University College London, Verenigd Koninkrijk*

CONTENTS

CHAPTER 1

General Introduction



1

AIM OF THIS THESIS

Have you ever been cross-country skiing? I vividly remember stepping on these dangerously narrow planks for the first time. It was a sunny winter afternoon in the February of 2017 when I first set foot onto the slopes of the Bymarka nature reserve near Trondheim, Norway. I recall that I was with my beautiful wife Anja when I was taking my first steps on the brand-new cross-country skis that we had bought the day before. Driven by athletic ambition, but unsure about where to find a beginner-friendly course, we followed a local on a slope into the woods. We were immediately faced with a descent, far too steep for my abilities, which hardly allowed me to stay upright and move forwards on my skis. I will leave to your imagination what happened next.

Not only do I remember the event of going cross-country skiing for the first time, but I also recall when and where it happened. Recollecting such a so-called episodic memory (Tulving, 1972, 2002) includes our subjective sense for where in space and when in time the remembered event occurred. Animal and human neuroscience research has elucidated many of the mechanisms that allow us to know where we are in space as well as when in time our experience unfolds. Specifically, the hippocampal-entorhinal region, situated deep in the medial temporal lobe of the human brain, has been scrutinized in search of spatial and temporal codes. In rodents, functionally defined cell types have been described that encode positions in space (Moser et al., 2017) and moments in time (Eichenbaum, 2014). Comparable spatial codes are thought to underlie spatial navigation in humans (Epstein et al., 2017). Studies of the human brain have particularly highlighted the contribution of the hippocampal-entorhinal region to episodic memory (Squire and Zola-Morgan, 1991; Moscovitch et al., 2016).

In this thesis, I address the question how coding principles of the hippocampal-entorhinal region support our memory for where and when events unfold. I aim to illuminate how we use coding mechanisms that have originally been discovered in research on how the mammalian brain represents space and time to organize our memory and experience. Thus, my thesis speaks to the overarching question how coding mechanisms in the hippocampal-entorhinal region provide a foundation for memory and contribute to cognitive functions more generally. To this end I will describe findings from three experiments: In a first study, I investigate how we use spatial codes to picture imagined scenarios in front of our mind's eye, speaking to the idea that these codes underlie our ability to draw on past experience to mentally

simulate future scenarios. Next, in a behavioral experiment, I consider the impact of distorted spatial codes on our memory for where events occurred. Subsequently, I turn towards our memory for when in time events take place. I examine how learning a sequence of events shapes our memories for these events in the entorhinal cortex. In addition to this empirical work, I draw upon a review of literature from both animal and human research to propose a theoretical account in which spatial processing principles are at the heart of a domain-general framework for the role of the hippocampal-entorhinal region in a wide range of cognitive functions. Together, the empirical and theoretical work I will present in this thesis aims to elucidate how hippocampal-entorhinal coding principles operate across information domains to organize our memory and experience.

Focusing on the hippocampal-entorhinal region, I will begin by introducing landmark findings from spatial navigation research to lay out how the brain processes and represents space. Further, I will complement the description of these coding principles with insights from temporal memory research. Additionally, I will provide background information on the anatomy of the hippocampal-entorhinal region and the experimental and analytical techniques used in the experiments described in this thesis.

THE REPRESENTATION OF SPACE IN THE HIPPOCAMPAL-ENTORHINAL REGION

A representation of our location in space is crucial for navigation. We need to know our location and the location of our goal to navigate in a goal-directed manner, for example when trying to find our way to a warm cabin to escape the cold on a long cross-country skiing tour. Spatial navigation has been studied extensively in animal models, in particular using electrophysiological recordings in freely moving rodents. In a typical experiment, an animal forages for food in a small enclosure while neuronal activity is recorded. In the human brain, navigation can be studied noninvasively using functional magnetic resonance imaging (fMRI) and virtual reality (VR) technology. In this section, I will describe some of the core findings these studies have yielded that shape our understanding of how space is represented in the mammalian brain in general and the hippocampal-entorhinal region in particular. Throughout this thesis I will explore how these spatial coding mechanisms support cognitive functions beyond navigation.

Place cells: A map of space in the hippocampus

During navigation place cells carry information about the animal's location in space. These cells, first described by O'Keefe and Dostrovsky (1971), fire only when the animal occupies a certain region in space, which is referred to as its place field. Different cells fire at different locations in the environment and collectively the firing fields of the population of place cells are thought to provide a map-like representation of the animal's surroundings (O'Keefe and Nadel, 1978). Place cells were first discovered in rats, but have since been observed in humans (Ekstrom et al., 2003; Miller et al., 2013; Qasim et al., 2018), mice (Rotenberg et al., 1996) and bats (Ulanovsky and Moss, 2007; Yartsev and Ulanovsky, 2013). Place cells signal the location of the animal in relation to external cues, which can control the location of place fields in the environment (O'Keefe and Conway, 1978; Muller and Kubie, 1987). Intriguingly, place cells 'remap' between environments, so that independent maps are formed for different environments (Muller and Kubie, 1987; Bostock et al., 1991; Leutgeb et al., 2004). The subsets of active place cells differ between environments and cells exhibiting place fields in multiple environments do so at locations that cannot be predicted from one environment to the next (Bostock et al., 1991; Leutgeb et al., 2004).

Since evidence for a topographic organization of place cells is lacking (O'Keefe et al., 1998; Redish et al., 2001), fMRI evidence for positional representations in the human brain (Hassabis et al., 2009; Rodriguez, 2010; Sulpizio et al., 2014; Kim et al., 2017) remains controversial as it is unclear how place cell activity could translate to location-specific differences in BOLD-activity (Nolan et al., 2018). While a topographic organization of place cells has not been observed (O'Keefe et al., 1998; Redish et al., 2001), the width of place fields exhibits a gradient along the dorso-ventral axis of the rodent hippocampus. On an 18m-long linear track, place field width increased from less than 1m at dorsal recording sites to around 10m in the ventral hippocampus (Kjelstrup et al., 2008). In *Chapter 5*, I review evidence for this anatomical gradient across species and cognitive functions and consider its implications for human cognition.

Grid cells: A coordinate system for the map

Whereas a place cell is typically characterized by bursts of activity constrained to one location in the environment, grid cells in the medial entorhinal cortex, one synapse upstream of the hippocampus (see Box 1), exhibit multiple firing fields. These firing fields are located at the vertices of equilateral triangles tiling the entire environment, thus resulting in a 60°-symmetric firing pattern (Hafting et al., 2005).

These firing patterns are described based on the size of the firing fields and the spacing between them as well the orientation and translational offset of the pattern relative to the enclosure (Hafting et al., 2005; Moser et al., 2014, 2017; Rowland et al., 2016). The regular firing patterns of grid cells are thought to provide a metric for space and support path integration (Hafting et al., 2005; McNaughton et al., 2006; Burak and Fiete, 2009). Path integration refers to the ability to keep track of one's location based on self-motion cues; allowing the animal to update its homing vector (Mittelstaedt and Mittelstaedt, 1980; Gallistel, 1990; Etienne and Jeffery, 2004).

Comparable to place field size, grid scale varies along the dorso-ventral axis of the medial entorhinal cortex with grid patterns of larger scale at more ventral recording sites (Hafting et al., 2005; Barry et al., 2007; Brun et al., 2008; Stensola et al., 2012). However, this increase in grid-scale is step-like rather than continuous, with cells sharing a similar scale and orientation clustering in so-called modules (Hafting et al., 2005; Barry et al., 2007; Brun et al., 2008; Stensola et al., 2012). In contrast to place cells, the firing properties of the entorhinal grid system are thought to be more rigid. The relationships between the firing fields of grid cells from the same module are consistent across environments with coherent changes of the offset and rotation of the grid pattern across cells from a module (Fyhn et al., 2007; Yoon et al., 2013).

Next to rats (Hafting et al., 2005), mice (Fyhn et al., 2008) and crawling bats (Yartsev et al., 2011), grid cells have been recorded from the entorhinal cortex and cingulate cortex of human epilepsy patients undergoing presurgical screenings (Jacobs et al., 2013; Nadasdy et al., 2017). In non-invasive investigations of grid-like coding in the healthy human brain, hexadirectional signals are considered as an fMRI-signature of population activity in the entorhinal grid system (Doeller et al., 2010). Here, BOLD-activity is analyzed as a function of running direction through a virtual enclosure. One partition of the dataset is used to estimate the orientation of the hexadirectional signal relative to the environment and, in an independent data partition, activity is subsequently examined as a function of running direction relative to this estimated orientation (Doeller et al., 2010). Using this technique, hexadirectional signals have been observed in the human entorhinal cortex and a network of regions implicated in episodic memory (Doeller et al., 2010). Recently, hexadirectional modulations of theta and broadband high-frequency power of MEG and intracranially recorded EEG data have been observed as a function of virtual running and saccade direction (Chen et al., 2018; Maidenbaum et al., 2018; Staudigl et al., 2018). Furthermore, the investigation of hexadirectional signals has

opened an avenue to study a potential mechanism underlying the breakdown of spatial navigation abilities in aging and disease. Here, compromised hexadirectional signals have been reported for healthy young adults carrying a genetic risk factor for Alzheimer's disease (Kunz et al., 2015) compared to non-carriers as well as for elderly compared to young adults (Stangl et al., 2018). One explanation for these findings is the reduced stability of the orientation of the hexadirectional signal over time, resulting in inconsistent orientations between data partitions (Kunz et al., 2015; Stangl et al., 2018).

Head direction cells: The map's compass

Next to self-location, information about one's orientation relative to relevant landmarks is crucial for successful navigation. This is thought to be encoded by head direction cells (Taube et al., 1990a, 1990b; Taube, 2007), which are active whenever the animal faces a certain direction irrespective of its position and thereby provide compass-like information (Taube et al., 1990a, 1990b). These cells have been discovered in the subiculum and have since been described throughout a distributed network including the anterior dorsal thalamic nucleus, the lateral mammillary nuclei, entorhinal and retrosplenial cortex (see Taube (2007) for review). The head direction signal is thought to be generated via the integration of vestibular and proprioceptive inputs (Taube, 2007; Cullen and Taube, 2017), but it is anchored to cues in the environment. For example, the preferred firing directions of head directions follow rotations of salient visual cues (Taube et al., 1990b).

In humans, directional representations have been demonstrated using fMRI. Repetition enhancement and repetition suppression effects have been interpreted as evidence for absolute directional representations in medial parietal areas, parahippocampal cortex and the thalamus (Baumann and Mattingley, 2010; Vass and Epstein, 2013; Shine et al., 2016). In these studies, visual stimuli cueing four different directions were presented, resulting in differential BOLD-responses for repetitions of the same direction compared to different directions in consecutive trials. Similarly, the BOLD-response was modulated by the time since the last trajectory in the current direction during virtual navigation, demonstrating the sensitivity of parahippocampal as well as retrosplenial and visual cortices to heading or view direction (Doeller et al., 2010). Consistently, directional coding has also been observed using multivariate analyses techniques. Here, cues for the same direction elicited more similar multi-voxel patterns in the presubiculum, entorhinal cortex and the retrosplenial complex than cues for different directions (Vass and Epstein, 2013, 2016).

A potpourri of functionally-defined cell types

Beyond place, grid and head direction cells, an increasing number of functionally defined cell types has been identified (for review see Moser et al. (2017); Grieves & Jeffery (2017); Poulter et al. (2018)). Among them are conjunctive grid×head direction cells, particularly in the deeper layers of the entorhinal cortex, which encode both position and head direction (Sargolini et al., 2006). Further, speed cells in the medial entorhinal cortex signal the animal's running speed. These cells are described by firing rates linearly covarying with running speed (Kropff et al., 2015). Border cells in the medial entorhinal cortex fire along the walls of the environment (Solstad et al., 2008; Savelli et al., 2008), whereas boundary vector cells in the subiculum are active only at specific distances and directions to these boundaries (Lever et al., 2009). Similarly, object vector cells have recently been described to encode distances and directions to objects in the environment (Høydal et al., 2018). In flying bats, vectors to navigational goals are encoded by cells sensitive to the egocentric direction and the distance to a goal (Sarel et al., 2017).

MEMORY ORGANIZATION AND THE HIPPOCAMPUS

While research on the rodent hippocampal formation has long focused on its role spatial navigation, the role of the hippocampus has been scrutinized for decades in human memory research. The description of the patient H.M. has had a long-standing impact on our understanding of human memory (Scoville and Milner, 1957; Squire, 2009). As a treatment for his epileptic seizures, large parts of H.M.'s medial temporal lobes, including bilateral hippocampus, were surgically removed. After the surgery, H.M. suffered from a severe anterograde amnesia, leaving him unable to form new memories for facts or events (Scoville and Milner, 1957). Further, H.M. was afflicted with a temporally graded retrograde amnesia, leaving him unable to remember information learned recently before his surgery, but able to recall earlier memories, for example from his childhood. Notably, some of H.M.'s mnemonic abilities appeared spared. In a mirror drawing task H.M. improved over several days of testing; despite having no recollection of having been tested on this task before (Milner, 1962; as cited in Squire, 2009).

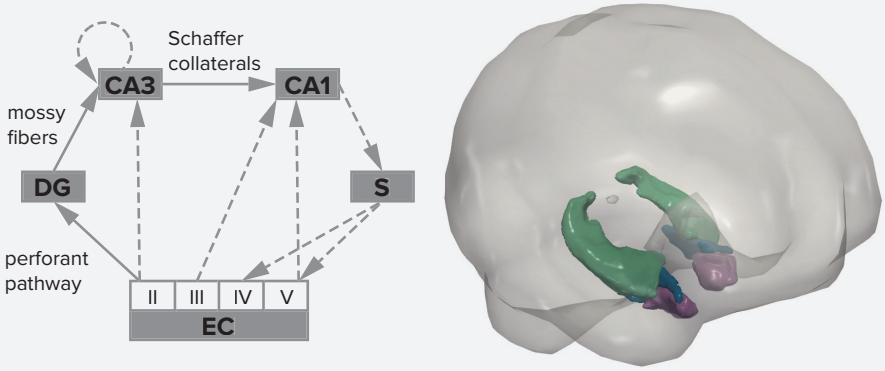
The retained ability of amnesic patients to acquire new skills provided first evidence for the notion of multiple, anatomically distinct memory systems in the brain (Cohen and Squire, 1980; Squire, 1982, 1992). In a classic taxonomy of different forms of memory, declarative (or explicit) and non-declarative (or implicit) memory are

distinguished (Squire, 1982, 1992; Squire and Zola-Morgan, 1991). Non-declarative memory was thought to include skills, habits, priming, classical conditioning and non-associative learning. The symptoms of amnesic patients suggested these abilities to be independent of the medial temporal lobe. In contrast, declarative memory appeared to crucially depend on medial temporal lobe structures, including most prominently the hippocampus. In this taxonomy, declarative memory was thought to be composed of episodic and semantic memory (Squire, 1982, 1992). Memories for specific events, such as one's first adventure on cross-country skis, that include information about the temporal and spatial context in which they occurred, are typically referred to as episodic memories. In contrast, memories for facts such as that Oslo is the capital of Norway, are usually defined as semantic memories (Tulving, 1972, 1985, 2002; Squire, 1982, 1992).

Focusing on the associative nature of memory, other accounts have emphasized the hippocampal contribution to memory via its role in the formation of associations between previously unrelated items (Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993; Eichenbaum et al., 1999; Eichenbaum and Cohen, 2014; Henke, 2010). In this framework, the hippocampus forms relational networks of arbitrary associations. Different items constitute nodes between which links are formed and different relational networks can be linked via overlapping nodes (Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993; Eichenbaum et al., 1999). Seminal work in rodents has demonstrated that hippocampal lesions disrupt performance on tasks requiring transitive inference or order reversals of paired elements, underscoring the central role of the hippocampus in the flexible expression of memories (Bunsey and Eichenbaum, 1996; Dusek and Eichenbaum, 1997).

Box 1: Anatomy of the hippocampal-entorhinal system

The hippocampal formation, consisting of the dentate gyrus (DG), the hippocampal subfields CA1, CA2, and CA3 as well the subiculum (S) and entorhinal cortex (EC), lies deep in the medial temporal lobe of the human brain. The hippocampus, which receives highly processed sensory inputs, is considered to constitute the apex of cortical processing streams (Felleman et al., 1991; Squire and Zola-Morgan, 1991) and its basic structure is relatively well-preserved across mammals (Clark and Squire, 2013; Witter et



al., 2017a). Notably, the hippocampal long-axis, which has a dorso-ventral orientation in rodents, has undergone a rotation of approximately 90° over the course of evolution so that the hippocampal longitudinal axis roughly corresponds to the anterior-posterior axis of the primate brain (Strange et al., 2014). In *Chapter 5*, I review evidence for how a functional gradient along the hippocampal long-axis might support cognitive processes. Classically, information flow through the hippocampus has been characterized by the trisynaptic loop (left panel, solid lines). Via the perforant pathway, Layer II of the entorhinal cortex projects to the dentate gyrus. These inputs are fed forward via the mossy fibers to CA3 and from there via the Schaffer collaterals to CA1 (van Strien et al., 2009; Clark and Squire, 2013; Witter et al., 2017a).

The entorhinal cortex in rodents is typically subdivided into a medial and a lateral subregion (Canto et al., 2008; Witter et al., 2017b). The functionally defined cell types encoding spatial information, including grid and head direction cells as well as speed and border cells have been identified mainly in the medial entorhinal cortex (Moser et al., 2017). Cells in the lateral entorhinal cortex have been shown to be selective for the identity of specific objects (Deshmukh and Knierim, 2011; Tsao et al., 2013; Knierim et al., 2014) and recently temporal information has been decoded from population activity in the lateral entorhinal cortex (Tsao et al., 2018). Next to diverging functional signatures, the two entorhinal subdivisions also exhibit different connectivity profiles. Capitalizing on differential functional connectivity fingerprints, two independent studies using high-field fMRI reported converging evidence for a bipartite division of the human entorhinal cortex

(Navarro Schröder et al., 2015; Maass et al., 2015) (right panel shows front-right view of a human brain with the hippocampus, posterior-medial and anterior-lateral entorhinal cortex in green, blue and purple, respectively). Navarro Schröder et al. (2015) focused on whole-brain connectivity patterns to describe both an anterior-posterior and a medial-lateral gradient in the connectivity profiles of entorhinal voxels, with stronger connectivity to medial-prefrontal and orbitofrontal regions for anterior-lateral entorhinal cortex in contrast to stronger connectivity to occipital and parietal regions for the posterior-medial entorhinal cortex. Maass, Berron et al. (2015) investigated functional connectivity within the medial temporal lobe. Consistent with the notion that the human homologue region of the rodent medial and lateral entorhinal cortex correspond to the posterior-medial and anterior-lateral human entorhinal cortex, their findings indicate stronger connectivity between the parahippocampal cortex and the posterior-medial entorhinal subregion, whereas the anterior-lateral subregion was more strongly connected to the perirhinal cortex (Maass et al., 2015). Focusing on hexadirectional coding in the posterior-medial and temporal mapping in the anterior-lateral entorhinal cortex, the work presented in *Chapters 2 and 4* sheds light onto the question how entorhinal subdivisions differentially contribute to human cognitive function.

The empirical work presented in this thesis addresses the question how the hippocampal-entorhinal region supports human episodic memory and related cognitive processes. *Chapter 2* speaks to the question how spatial coding principles support flexible memory expression in the form of mental simulation. Returning to the notion that episodic memory encompasses knowledge about where in space and when in time events occurred, I investigate memory for spatial and temporal information in *Chapters 3 and 4*. I assess in a behavioral experiment in *Chapter 3* whether the accuracy of spatial memory can be influenced by the geometry of the environment in which memories are formed. In *Chapter 4*, I assess how event representations in the entorhinal cortex are shaped by learning about when in time these events occur. I focus on these aspects as examples for key coding principles by which processing mechanisms in the hippocampal-entorhinal region could support human memory and cognitive functions more broadly.

SPATIAL CODES FOR MEMORY AND IMAGINATION

What is the role of spatially tuned cells and, more generally, spatial codes in the hippocampal-entorhinal region in guiding cognition and behavior? In this thesis, I address these long-standing questions empirically by focusing on spatial memory and mental simulation in *Chapters 2 and 3*. More broadly, a central role for spatial computations across cognitive domains is at the heart of the theoretical account I outline in *Chapter 5* based on the conjunction of recent findings about non-spatial coding of place and grid cells with a spatial representational format for cognition.

Grid cells as a spatial metric for memory

Ever since their discovery, grid cells have been suggested to support path integration (Hafting et al., 2005; McNaughton et al., 2006; Burak and Fiete, 2009). While a direct demonstration of the involvement of grid-cell computations in homing behavior in rodents remains elusive, recent findings point towards impaired path integration in mice with disrupted grid cell firing (Gil et al., 2018). Here, grid cell firing was abolished by removing NMDA glutamate receptors from the medial entorhinal cortex and subicular regions. Compared to control mice, these animals exhibited increased angular errors in their swimming paths to an occluded platform they had previously been guided to by swimming along a corridor including a turn (Gil et al., 2018). In humans, evidence for the importance of the entorhinal grid system for path integration stems from a task in which participants estimated distances and directions to their start point from different positions along a guided path (Stangl et al., 2018). In older, but not younger adults, weaker hexadirectional signals in the entorhinal cortex correlated with increased path integration errors (Stangl et al., 2018). Taken together, these findings suggest the integrity of the entorhinal grid system to be important for the calculation of vectors to target locations based on the integration of self-motion.

Finding trajectories between locations requires a representation of the displacement between them. The regular firing patterns of grid cells are thought to encode positions in the environment. While the firing pattern of each cell is periodic and hence provides only ambiguous spatial information, the combination of spatial phases across grid modules with different scales provides an efficient positional code (Fiete et al., 2008; Mathis et al., 2012; Bush et al., 2015; Carpenter and Barry, 2016; Herz et al., 2017). Positions encoded by this grid cell population phase can be used to derive vectors between positions to guide navigation (Bush et al., 2015; Carpenter and Barry, 2016). Indeed, an artificial agent endowed with grid-like representations

of its surroundings exhibited the ability to use these for vector navigation (Banino et al., 2018). In *Chapter 2*, I build on theoretical work showing that the entorhinal grid system can be used to compute vectors between positions stored in memory and test for a 60°-modulation of entorhinal multi-voxel patterns while participants retrieved the locations of start and goal positions from memory to imagine spatial views from the start location facing towards the goal in a virtual city.

Underscoring the role of the entorhinal grid system in memory for locations, hexadirectional signals in the entorhinal cortex have been linked to the accuracy of spatial memory. In across-subjects analyses, both the coherence of estimated hexadirectional orientations across voxels (Doeller et al., 2010) and the strength of the hexadirectional signals (Kunz et al., 2015) predicted memory for object positions in the testing environment. Moreover, assuming hexadirectional signals are indeed driven by grid cell activity, these findings suggest the regularity of grid firing patterns to be central for spatial memory performance; in line with theoretical work capitalizing on the properties of grid firing patterns to demonstrate that positions can be stored using grid codes across modules (Fiete et al., 2008; Herz et al., 2017; Mathis et al., 2012) and that these can be used to guide navigation behavior (Bush et al., 2015; Carpenter and Barry, 2016; Banino et al., 2018).

Box 2: Virtual reality

During navigation the vestibular system provides information about rotations and translations of the head and proprioceptive sensors signal the position of limbs. Furthermore, movement results in optic flow perceived via the visual system. While navigation is frequently studied using electrophysiological recordings in freely moving rodents, engaging all systems, human cognitive neuroscience typically relies on virtual reality technology to study the neural systems underlying navigation using neuroimaging (Maguire et al., 1998; Hartley et al., 2003; Wolbers et al., 2004, 2007; Wolbers and Büchel, 2005; Spiers and Maguire, 2006, 2007; Doeller and Burgess, 2008; Doeller et al., 2008, 2010). Only the use of virtual environments, which participants can navigate using joysticks or key presses, allows the concurrent recording of brain activity since both fMRI and MEG require the minimization of physical movements. This reduces the information content compared to physical navigation with the visual system providing the dominant input. Hence, virtual

navigation in fMRI, where participants assume a supine position, has been argued to fail to capture the same neural processes as physical navigation (Taube et al., 2013).

While virtual navigation inarguably does not engage proprioceptive, vestibular and motor systems in the same way as physical navigation, neuroimaging studies have shed light onto the question how the human brain supports navigation. Early studies contrasted for example periods of navigation versus rest and navigational demands between wayfinding and route following to provide evidence for the involvement of the hippocampus in human navigation (Maguire et al., 1998; Hartley et al., 2003). More specifically, studies using virtual environments demonstrated coding of facing directions (Baumann and Mattingley, 2010; Shine et al., 2016), directions to goal (Chadwick et al., 2015), directional heading in local reference frames (Marchette et al., 2014) and hexadirectional signals (Doeller et al., 2010; Kunz et al., 2015; Horner et al., 2016; Navarro Schröder et al., 2017; Stangl et al., 2017, 2018; Maidenbaum et al., 2018; Chen et al., 2018). In *Chapter 2*, I build on these findings to investigate absolute and hexadirectional coding during imagination of views in a virtual city.

Recent developments in virtual reality technology open up exciting avenues to enhance the feeling of immersion, the perception of being in the virtual environment. These advances include the use of head mounted displays, which combine the stereoscopic presentation of images with wide field of view lenses to create the perception of three-dimensional space as well as the tracking of head position to update visual inputs based on movements. Especially in combination with novel motion tracking platforms, which translate physical steps and rotations into virtual movements, virtual environments become highly immersive and navigation through them engages motor, proprioceptive and vestibular systems as well. I harness the enhanced immersion provided by a head mounted display and a motion platform to study the impact of environmental geometry on spatial cognition in *Chapter 3*. Advances in the development of portable MEG sensors (Boto et al., 2018) bear exciting potential in that they might soon enable the study of neural signals during navigation using highly immersive virtual reality setups.

Intriguingly, virtual reality also allows researchers to manipulate the subject's experience in ways impossible in the real world. For example, alterations of rotational or translational gain have been used to decouple contributions of visual and interoceptive inputs to human path integration (Tcheang et al., 2011) and rodent place (Chen et al., 2013) and grid cell firing (Campbell et al., 2018). Further, the field of view of participants can been tightly controlled (Auger et al., 2015) or spatial coding in the vertical dimension during volumetric trajectories can be investigated (Kim et al., 2017; Kim and Maguire, 2018). In *Chapter 4*, I leverage the unique potential of virtual reality by introducing teleporters along the route through a virtual city; thereby disentangling the otherwise inevitably correlated dimensions of space and time.

However, a growing body of evidence demonstrates that grid cell firing can deviate from the canonical hexagonal pattern. Upon first encounter of a novel environment, grid firing patterns expand, resulting in reduced hexagonal symmetry (Barry et al., 2012). These expansions of the grid pattern are reversible and grid patterns return to their standard scale as the environment becomes familiar to the animal (Barry et al., 2012). Varying the size and shape of familiar rectangular enclosures, grid firing patterns were shown to directly follow expansions and compressions of environmental boundaries (Barry et al., 2007). Likewise, human path integration performance is influenced by changes to the aspect ratio of a rectangular enclosure (Chen et al., 2015).

Importantly, environmental boundaries can influence grid cell firing patterns also with prolonged experience. In familiar enclosures, the grid pattern is characterized by an elliptical distortion, suggested to arise from anchoring of the grid to reference points (Stensola et al., 2015). More drastically, degradations of the grid pattern were observed in a trapezoidal enclosure (Krupic et al., 2015). Compared to a square control environment, the grid firing fields were larger and the hexagonality of firing patterns was reduced. Contrasting the grid distortions between the two halves of this environment with highly-polarized geometry revealed most severe distortions of grid patterns in the narrow part of the trapezoidal enclosure: Overall hexagonal symmetry of the firing patterns was reduced in the narrow compared to the broad part of the trapezoid and grid patterns in the narrow end were characterized by larger firing fields and larger spacings between fields. Further, the orientation of

the grid pattern differed between the two halves of the trapezoid (Krupic et al., 2015). Consistently, recent findings corroborate the notion that environmental geometry can distort grid firing patterns (Krupic et al., 2018). In a quadrilateral environment, in which the angle of one slanted wall was varied across trials, grid firing patterns were shifted away from the slanted wall. In a simulation capturing this local distortion of the grid pattern, positional decoding was less accurate near the slanted wall than further away from it (Krupic et al., 2018). Taken together, these findings demonstrate that environmental geometry exerts a strong influence on grid firing patterns; resulting in grid pattern distortions in environments with highly polarized geometry. Further, they stand in contrast to the canonical regular grid patterns assumed by models of grid cell function.

In *Chapter 3*, I address the open question which consequences might arise from distorted spatial codes for our memory of where in space events occur. I focus on the effects of putative grid pattern distortions in a trapezoidal environment on spatial cognition. Following the assumption that locations are encoded in grid cell population phases and that vectors between positions can be calculated from these representations (Fiete et al., 2008; Mathis et al., 2012; Bush et al., 2015; Herz et al., 2017; Banino et al., 2018), behavioral biases might result from grid pattern distortions (Carpenter and Barry, 2016). Hence, studying spatial cognition in environments with highly polarized geometry, in which the regularity of entorhinal grid patterns is strongly degraded in rodents, opens up an opportunity to study how the entorhinal grid system supports spatial memory. In *Chapter 3,* I use immersive virtual reality technology (see Box 2) to investigate how spatial cognition is impacted by environmental geometry — both within and outside of a trapezoidal environment.

Spatial building blocks for mental simulation

The retrieval of episodic memories has been conceptualized as mental time travel to the place and time where an event occurred (Tulving, 1983, 2002). Connecting the putative role of place and grid cells in storing positions and the notion that episodic simulations are facilitated by the reinstatement of contextual information, suggests a role for spatial representations in mental simulation. For example, when reminiscing about my first cross-country ski adventure, I might mentally travel back to the slopes of the Bymarka, putatively drawing upon similar representations of these surroundings as when actually learning to ski. Indeed, spatially tuned cells have been ascribed central roles in spatial memory retrieval and mental simulation more generally (Byrne et al., 2007; Buckner, 2010; Bicanski and Burgess, 2018). Place, grid and head direction cells have been proposed to underlie human recollection

and imagination when mentally simulating spatial views (Bicanski and Burgess, 2018; Byrne et al., 2007). In this framework, place cells in the hippocampus are thought to provide stored positional information with respect to external landmarks during episodic simulations, which can be combined with information about the location and identity of objects in the proximity. Building upon head direction cell firing, these allocentric representations can then be transformed to simulate views from an egocentric perspective (Bicanski and Burgess, 2018; Byrne et al., 2007). Grid cells are incorporated in this model as a means to shift the simulated viewpoint by updating the place cell representation (Bicanski and Burgess, 2018).

Intriguingly, this proposed role of spatial representations in human imagination is consistent with the structured, non-local activity of spatially tuned cells observed in rodents during so-called replay (see Foster (2017); Ólafsdóttir et al. (2018) for review). Here, hippocampal place cells fire in sequences of the same order as during navigation on a maze (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996). This phenomenon was first observed during sharp-wave ripple events during sleep, but also occurs during awake rest (Kudrimoti et al., 1999; Foster and Wilson, 2006; Diba and Buzsáki, 2007; Karlsson and Frank, 2009) and exploratory behavior (O'Neill et al., 2006). During replay, place cells maintain their spatial relationships to each other in that cells with overlapping place fields are co-active during sleep (Wilson and McNaughton, 1994) and the timing of activity reflects the distance between the cells' place fields (Skaggs and McNaughton, 1996; Diba and Buzsáki, 2007). These temporally compressed sequential patterns of activity (Skaggs and McNaughton, 1996; Lee and Wilson, 2002), which occur both in forward and reverse order (Diba and Buzsáki, 2007; Davidson et al., 2009), demonstrate the striking ability of place cells to represent remote locations and trajectories (Carr et al., 2011; Foster, 2017; Ólafsdóttir et al., 2018).

Is the replay of non-local positions restricted to place cells or could other functionally defined cell types contribute to this phenomenon as well? Evidence suggests grid cell firing to be consistent with trajectories replayed by place cell activity during rest and periods of task disengagement (Ólafsdóttir et al., 2016, 2017). During rest, the firing of grid cells in the deep layers of the entorhinal cortex, which receive input from the hippocampus, was increased during replay events defined by hippocampal place cell firing. Importantly, grid cells fired in spatial coherence with the trajectories replayed by place cell activity (Ólafsdóttir et al., 2016). However, grid cell replay in the superficial layers of the entorhinal cortex appears to be independent of hippocampal place cells (O'Neill et al., 2017) and the precise interactions between

place and grid cell firing during replay are not fully understood. Persistent spike-timing correlations between grid cell firing during spatial exploration and during sleep further illustrate the maintenance of firing relationships in the entorhinal grid system across behavioral states (Gardner et al., 2017; Trettel et al., 2017). Similarly, during sleep, head direction cells fire in coherence with offsets of their preferred orientations observed during navigation (Peyrache et al., 2015; Gardner et al., 2017). Taken together, these findings suggest spatial coding properties of grid and head direction cells to persist across behavioral states; opening up the possibility for coherent activity of spatially tuned cells during replay.

Place cell replay has been suggested as a potential neural substrate for the role of the hippocampus in future anticipation, imagination and planning (Buckner, 2010). Here, replay is conceptualized as a reflection of event sequences (Buckner, 2010). Focusing on the predictive function of memory, hippocampal sequences might allow the simulation of past and future events and the flexible recombination of different elements to construct novel simulations (Buckner, 2010). These ideas are in line with the central role of the hippocampus in human imagination (Buckner and Carroll, 2007; Hassabis and Maguire, 2007; Hassabis et al., 2007a, 2007b; Schacter et al., 2007, 2012). Neuroimaging experiments where participants are cued to generate fictitious scenarios demonstrate the activation of the hippocampus as a key component of a core system of brain regions thought to underlie episodic simulations (Hassabis and Maguire, 2007; Hassabis et al., 2007b; Schacter et al., 2007, 2012; Spreng et al., 2009; Benoit and Schacter, 2015). Further, the ability to imagine novel scenarios is impaired in patients with hippocampal lesions (Hassabis et al., 2007a). The debate on the precise role of the hippocampus in imagination is ongoing with prominent accounts suggesting it to enable the flexible recombination of episodic details (Schacter et al., 2007, 2012), while others emphasize its role to be in the construction of coherent scenes when picturing novel scenarios in front of the mind's eye (Hassabis and Maguire, 2007).

The recombination of past experience to construct potential future scenarios is a core adaptive function of flexible memory expression. Spatial codes in the hippocampal-entorhinal region might provide the building blocks for these simulations, for example in service of navigational planning. To address the question which role spatial representations play for mental simulation, I built upon the established role of the medial temporal lobe in imagination in general and the suggested contributions of spatially tuned cells to mental imagery specifically. I tested for representations of absolute facing direction and a 60°-modulation of entorhinal pattern similarity

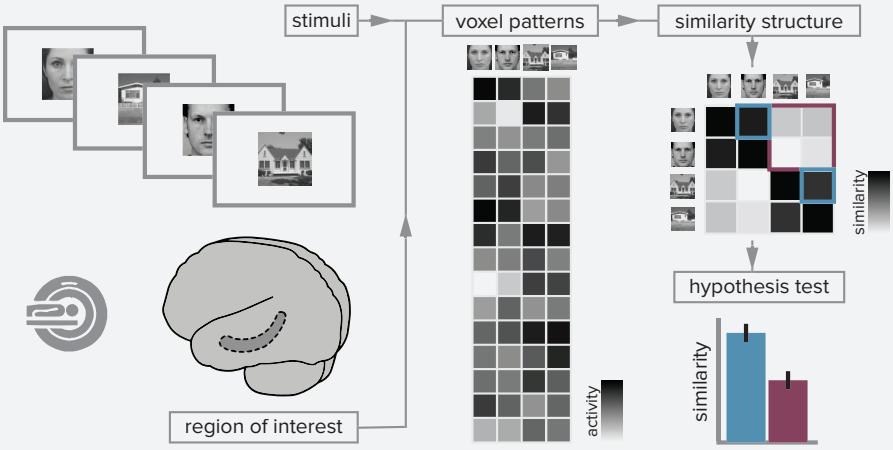
in an imagination paradigm in *Chapter 2* (see Box 3). In this task, participants imagined spatial views based on the learned positions of buildings in an urban virtual environment. I hypothesized representations of the absolute facing direction to be evident in increased pattern similarity when participants imagined similar directions, consistent with a population response of the head direction system. Further, I predicted activity of the entorhinal grid system, in line with its proposed role in the calculation of vectors from start to goal positions (Bush et al., 2015) and the updating of imagined viewpoints (Bicanski and Burgess, 2018), to be reflected in the structure of entorhinal multi-voxel pattern similarity.

TEMPORAL CODING IN THE HIPPOCAMPAL-ENTORHINAL REGION

Next to where in space, when in time and in which order events occur constitutes a central aspect of our experience. I not only remember where I had my first experience on cross-country skis, but also when this event took place and what happened before (taking a crowded bus to get to the slopes) and after (eating a waffle with brown cheese). Indeed, temporal information about when and in which order experience unfolded is considered a fundamental pillar of episodic memory (Tulving, 1983, 2002; Eichenbaum, 2014; Ekstrom and Ranganath, 2017; Ranganath, 2018; Howard, 2018). How does the hippocampal-entorhinal region represent temporal relationships between different events and memories?

The hippocampal-entorhinal region is sensitive to the order in which events are experienced. For example, hippocampal activity increases in response to violations of expected stimulus order, in line with a mismatch signal for unexpected stimulus sequences (Kumaran and Maguire, 2006; Barnett et al., 2014) and multivariate hippocampal activity patterns carry information about the conjunction of stimulus identity and position in learned sequences (Hsieh et al., 2014). This is further corroborated by findings that hippocampal and entorhinal activity reflects associations between stimuli, which were implicitly learned from probabilistic stimulus sequences derived from underlying graph structures (Schapiro et al., 2012; Garvert et al., 2017). Furthermore, during learning of noun-triplets, greater activations in the bilateral hippocampus predict correct later recall of triplet order, but not triplet recognition, highlighting the role of the hippocampus in the encoding of temporal order information (Tubridy and Davachi, 2011). Analyzing hippocampal multi-voxel patterns while participants made semantic judgments about sequentially

Box 3: Representational Similarity Analysis



Functional magnetic resonance imaging capitalizes on the different magnetic properties of oxygenated and deoxygenated hemoglobin. The flow of oxygenated blood is thought to increase in areas of elevated neuronal activity. This results in a net decrease of deoxygenated blood near areas of neural activity. Since deoxygenated hemoglobin interferes more with the transverse magnetization of nuclei than oxygenated hemoglobin, magnetic resonance sequences sensitive to T2* decay measure stronger signals where blood carries relatively more oxygenated hemoglobin (Huettel et al., 2004). Though the precise link between neuronal activity and the so-called blood-oxygenation-level dependent (BOLD) response is not fully understood, BOLD-fMRI is considered an indirect measure of neural activity (Huettel et al., 2004; Logothetis, 2008). Most studies investigate the magnitude of the BOLD-response, for example in the subsequent memory paradigm brain activity at encoding is contrasted between stimuli based on whether they are later remembered or forgotten to investigate brain areas relevant for memory formation (Brewer et al., 1998; Wagner et al., 1998; Fernández et al., 1999). Typically, general linear models are used to quantify the fit between the activity time course of each voxel and events defined by the experimental design (Friston et al., 1994). This mass-univariate approach has yielded important insights into brain function across cognitive domains.

However, the mass-univariate approach disregards information carried by activity patterns across voxels. In contrast, multi-voxel pattern analysis (MVPA) examines the relationship between the activity levels of a set of voxels to capitalize on information in distributed activity patterns (Haxby et al., 2001; Haynes and Rees, 2005; Kamitani and Tong, 2005; Polyn et al., 2005; Kriegeskorte et al., 2006; Norman et al., 2006). In this thesis, I employ a particular form of MVPA, representational similarity analysis (RSA, see figure), which examines the relationships of multi-voxel patterns to investigate their representational content (Kriegeskorte et al., 2006, 2008a). Between different stimuli or experimental conditions (top left), patterns of activity (center) are compared across all voxels in an anatomically or functionally defined region of interest (bottom left) or, in a so-called searchlight analysis, in local, typically spherical, subsets of voxels (Kriegeskorte et al., 2006). These local search spheres are then iteratively moved throughout the entire brain. The resulting correlations between multi-voxel patterns (top right) are assessed based on hypotheses about the pattern similarity structure (bottom right). Using the searchlight approach, I test in *Chapter 2* which brain regions exhibit more similar multi-voxel patterns when participants imagine the same compared to different directions, which might reflect population activity of the head direction cell system. Further, I employ RSA in region of interest analyses to investigate the pattern similarity structure in the entorhinal subregions in *Chapters 2 and 4*.

presented object images, Jenkins and Ranganath (2016) demonstrate that less similar hippocampal activity patterns at encoding were correlated with accurate recall of temporal order information. Consistent with the hippocampal involvement in the encoding of stimulus sequences, the hippocampus is also activated during the retrieval of order information (Copara et al., 2014). Taken together, these findings suggest that temporal order information is central to mnemonic processing of event sequences in the hippocampus, letting me remember the order of events on the day of my first cross-country skiing trip.

Does the hippocampal-entorhinal region also process temporal relationships between events beyond the order in which they occur? In a mismatch detection paradigm, hippocampal pattern similarity between study and test sequences was increased for sequences with identical inter-stimulus intervals compared to sequences in which the

order of stimuli was the same, but the intervals between stimuli were varied with respect to the learned sequence (Thavabalasingam et al., 2018). These findings suggest that duration information might be integrated into hippocampal sequence encoding and dovetail with evidence that hippocampal activity at encoding predicts later temporal memory. In one study, images were presented in sequences of four during fMRI scanning and in a later surprise memory test participants were to indicate when over the course of the experiment they had encountered each object. Hippocampal activity following sequence encoding was higher for more compared to less accurate later memory for the temporal positions of individual images (Jenkins and Ranganath, 2010). A putative mechanism underlying memory for temporal relationships between events is provided by the decorrelation of hippocampal activity patterns (Ezzyat and Davachi, 2014; Kyle et al., 2015; Nielson et al., 2015; Deuker et al., 2016). For example, higher pattern similarity has been observed between stimuli remembered to be close together compared to stimuli remembered to be far apart, despite the same time having elapsed between their presentation (Ezzyat and Davachi, 2014). Likewise, hippocampal activity patterns correlate with the temporal distances between real-life events (Nielson et al., 2015) and pattern similarity change reflects the remembered temporal and spatial relationships between events encountered along a route through a virtual city (Deuker et al., 2016). Collectively, these findings strongly implicate the hippocampus in temporal memory and suggest that it might encode temporal relationships beyond stimulus order.

However, it remains unclear how information about the intervals between events arises in the hippocampus. One possibility is that temporal relationships are represented already in the entorhinal cortex, which provides important cortical input to the hippocampus (van Strien et al., 2009; Witter et al., 2017a, 2017b) (see Box 1). Indeed, one study demonstrates that entorhinal pattern similarity during encoding scales with later temporal memory. Participants listened to a radio story while undergoing fMRI and entorhinal pattern similarity during this encoding phase correlated with the later remembered length of the temporal intervals between events (Lositsky et al., 2016). However, the role of the entorhinal cortex in representing the temporal structure of memories during memory retrieval is underexplored and the more precise anatomical localization of temporal codes in the entorhinal cortex constitutes an open question.

In rodents, temporal coding has been investigated in the hippocampal-entorhinal region as well. Hippocampal time cells fire at certain time points during a temporal delay (Pastalkova et al., 2008; MacDonald et al., 2011; Eichenbaum, 2014; Mau et

al., 2018) and overlap with the population of place and grid cells during navigation (Kraus et al., 2013, 2015). Time cells have so far only been demonstrated during short intervals that the animal experienced repeatedly, leaving open the question how memory for longer temporal intervals might arise. In contrast, one recent study investigated temporal coding in the rodent lateral entorhinal cortex at timescales ranging from seconds to days (Tsao et al., 2018). Here, cells in the lateral entorhinal cortex exhibited a ramp-like code for time, characterized by firing rates linearly increasing or decreasing over time. Importantly, temporal information could be decoded from entorhinal population activity with high accuracy even if these cells were removed from the analysis, highlighting that cell populations collectively carry information about time at multiple scales. This temporal information is suggested to arise implicitly from the integration of unique experience (Tsao et al., 2018). Evidence for this notion stems from the observation that when the experience of the animal was constrained by training it to run repeated laps on a maze population activity in the lateral entorhinal cortex was more stable and traversed similar trajectories through its states of activity. Consequently, temporal decoding was improved for time within a lap, but global temporal coding was degraded compared to episodes of free foraging (Tsao et al., 2018). In line with these findings one recent study has pointed to the anterior-lateral subregion of the entorhinal cortex being involved in the retrieval of temporal information (Montchal et al., 2019). Here, participants had to recall when in time a scene from a sitcom had appeared during the episode watched in the experiment. BOLD-activity in the anterior-lateral entorhinal cortex as well as the perirhinal cortex and a network of brain regions including the hippocampus, the medial prefrontal cortex, posterior cingulate cortex and angular gyrus was larger for the third of trials in which memory responses were most precise compared to the third of trials where temporal memory was least accurate (Montchal et al., 2019).

Taken together, these findings suggest a differential role of the entorhinal subregions in processing temporal information for episodic memory. However, how memories are shaped by their underlying temporal structure remains unclear. In *Chapter 4*, I address the question how time organizes our memories in the entorhinal cortex. I ask how learning a temporal event structure influences the mnemonic representations of these events. Specifically, I focus on the anterior-lateral subregion of the entorhinal cortex to test the hypothesis that the anterior-lateral entorhinal cortex maps the temporal structure of experience to shed light onto the question how event representations are shaped by their temporal relationships. Further, I will speak to the question how entorhinal representations reflecting temporal structure relate to memory recall.

THESIS OUTLINE

Following this general introduction, I will empirically address the question how coding principles of the hippocampal-entorhinal region support mnemonic functions. Specifically, I will present three studies elucidating the mechanisms underlying human imagination as well as spatial and temporal memory, which tested hypotheses inspired from descriptions of coding principles in the hippocampal-entorhinal region in rodents. Further, I will present an idea for how the hippocampal-entorhinal region could contribute to cognitive functions in a domain-general framework.

In *Chapter 2*, I connect the involvement of the human medial temporal lobe in imagination (Buckner and Carroll, 2007; Hassabis and Maguire, 2007; Hassabis et al., 2007a, 2007b; Schacter et al., 2007, 2012) with findings of non-local activity sequences of spatially tuned cells in rodents (Foster, 2017; Ólafsdóttir et al., 2018) and theoretical work positing a central role for spatial codes in mental simulation and navigational planning (Byrne et al., 2007; Bush et al., 2015; Bicanski and Burgess, 2018; Banino et al., 2018). I test the idea that spatial codes in the medial temporal lobe are implicated in drawing on past experience to mentally travel to locations stored in memory and to imagine spatial views from there. I focus on directional representations in imagination to test for fine-grained absolute directional coding, putatively going back to population activity of the head direction system (Taube et al., 1990a, 1990b; Taube, 2007), and a hexadirectional modulation of multi-voxel patterns in the posterior-medial entorhinal cortex as a proxy measure for grid cell computations.

Building upon the striking distortions of rodent grid firing patterns in a trapezoidal enclosure (Krupic et al., 2015) and models of grid cell function in navigation and memory (Bush et al., 2015; Carpenter and Barry, 2016), I explore the role of environmental geometry in human spatial cognition in a behavioral experiment described in *Chapter 3*. This study speaks to the question whether spatial codes providing a metric for memory — putatively supported by the entorhinal grid system — might be subject to distortions under some circumstances. Using highly immersive virtual reality technology, I address the question if memory can be deformed by environmental geometry akin to how grid firing patterns in rodents are distorted in highly polarized environments and whether these mnemonic distortions persist outside of the encoding environment.

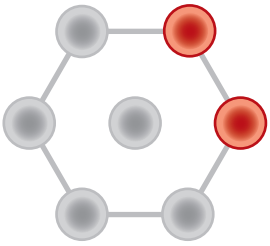
In *Chapter 4*, I turn towards the question how time, a core dimension of experience, organizes mnemonic representations in the human entorhinal cortex. Capitalizing on recent insight into temporal coding in the rodent lateral entorhinal cortex (Tsao et al., 2018) and evidence for a role of its human homologue region in temporal memory retrieval (Montchal et al., 2019), I assess how learning a temporal structure of events shapes event representations in the subregions of the human entorhinal cortex.

Following the experimental chapters, I present in *Chapter 5* a theoretical framework for how spatial processing principles in the hippocampal-entorhinal region could operate across domains of experience by providing a geometric code for high-level cognition. These ideas build upon the notion of a spatial representational format for cognition (Gärdenfors, 2000) and the review of a growing body of evidence implicating spatially tuned cells in the mapping of dimensions of experience beyond space for navigation. Specifically, I propose the continuous population code of place and grid cells to map the geometrically constrained dimensions of cognitive spaces. Stimuli are placed in these cognitive spaces based on their features along the dimensions spanning the space. The hippocampal-entorhinal region maps cognitive spaces at different levels of granularity, allowing both the retention of specific details and the generalization across experience. I suggest place cell remapping to enable the flexible formation of cognitive spaces with the more rigid firing properties of the entorhinal grid system providing a distance function for cognitive spaces. Further, temporally compressed sequences might enable the simulations of trajectories through cognitive space for optimal decisions.

Finally, in *Chapter 6*, I summarize the findings presented in *Chapters 2-4* and discuss them in light of the ideas presented in *Chapter 5*. In this general discussion, I consider the notion of cognitive spaces in the context of the cognitive map theory (Tolman, 1948; O'Keefe and Nadel, 1978) and the relational memory theory (Eichenbaum and Cohen, 1988; Eichenbaum et al., 1999; Eichenbaum and Cohen, 2014), two seminal accounts of how hippocampal computations use past experience to shape cognitive functions such as navigation and memory. Lastly, I outline outstanding questions for future research to elucidate how cognitive spaces in the hippocampal-entorhinal region might be supported by and interact with processing in other brain regions.

CHAPTER 2

Grid-cell representations in mental simulation



This chapter is published as:

J. L. S. Bellmund, L. Deuker, T. Navarro Schröder, C. F. Doeller (2016). Grid-cell representations in mental simulation. *eLife, 5*, e17089. https://doi.org/10.7554/eLife.17089

ABSTRACT

Anticipating the future is a key motif of the brain, possibly supported by mental simulation of upcoming events. Rodent single-cell recordings suggest the ability of spatially tuned cells to represent subsequent locations. Grid-like representations have been observed in the human entorhinal cortex during virtual and imagined navigation. However, hitherto it remains unknown if grid-like representations contribute to mental simulation in the absence of imagined movement. Participants imagined directions between building locations in a large-scale virtual-reality city while undergoing fMRI without re-exposure to the environment. Using multi-voxel pattern analysis, we provide evidence for representations of absolute imagined direction at a resolution of 30° in the parahippocampal gyrus, consistent with the head-direction system. Furthermore, we capitalize on the six-fold rotational symmetry of grid-cell firing to demonstrate a 60°-periodic pattern-similarity structure in the entorhinal cortex. Our findings imply a role of the entorhinal grid-system in mental simulation and future thinking beyond spatial navigation.

INTRODUCTION

Anticipation of the future is a central adaptive function of the brain and enables adequate decision-making and planning. Simulating or imagining future events and scenarios relies on a network of brain regions known to be involved in episodic memory, navigation and prediction (Buckner, 2010; Byrne et al., 2007; Hassabis and Maguire, 2007; Hasselmo, 2009; Schacter et al., 2012). For instance, before leaving your favorite cafe, you may picture the scenery in front of the cafe in your mind's eye to determine whether to take a left or a right turn to get home. To accomplish this you have to recall both the location of the cafe as well as the direction you are facing when leaving the building.

Electrophysiological recordings in freely moving rodents have demonstrated that positional information during navigation is represented by place cells in the hippocampus (O'Keefe and Dostrovsky, 1971) and grid cells in entorhinal cortex (Hafting et al., 2005). Place cells typically exhibit one firing field (O'Keefe and Dostrovsky, 1971), while grid cells are characterized by multiple firing fields arranged in a regular hexagonal pattern tessellating the entire environment (Hafting et al., 2005). Complementarily, directional information is carried by head direction cells, which increase their firing rate as a function of the animal's directional heading irrespective of its location (Taube et al., 1990a; Taube, 2007). Intracranial recordings in patients exploring virtual-reality (VR) environments demonstrated the existence of place and grid cells in the human hippocampus and entorhinal cortex, respectively (Ekstrom et al., 2003; Jacobs et al., 2010, 2013). A 60° directional periodicity of BOLD-signal modulations in the entorhinal cortex during virtual navigation indicates that grid-like entorhinal signals can also be detected with fMRI (Doeller et al., 2010; Kunz et al., 2015; Horner et al., 2016).

Notably, place cell activity can also represent locations other than the one currently occupied by the animal as illustrated by activation sequences corresponding to upcoming trajectories during rest periods (Dragoi and Tonegawa, 2011). Intriguingly, these 'preplay' sequences preferentially represent paths leading up to motivationally relevant locations (Ólafsdóttir et al., 2015). These observations support the notion that prospective coding of hippocampal place cells relates to the well-established role of the human hippocampus in mental simulation and imagination (Buckner, 2010; Byrne et al., 2007; Hassabis and Maguire, 2007; Hasselmo, 2009; Schacter et al., 2012). Akin to firing rate increases of neurons in the human medial temporal lobe specific to the content of imagination (Kreiman et al., 2000), firing patterns of

spatially tuned cells might be reinstated to imagine the view from a certain location during mental simulation (Bird et al., 2012; Byrne et al., 2007; Hasselmo, 2009). Prospective coding properties of grid cells (De Almeida et al., 2012; Kropff et al., 2015) and recent evidence for spatial coherence of grid with place cell activity during replay (Ólafsdóttir et al., 2016) further suggest a similar involvement of the entorhinal grid system in future anticipation and prediction. This is in line with the observation of grid-like representations during imagined movement through an environment (Horner et al., 2016). However, hitherto it remains unknown if grid-like representations support mental simulation independent of imagined movement, which could suggest a more general role of grid cell computations in navigational planning, future anticipation and cognition.

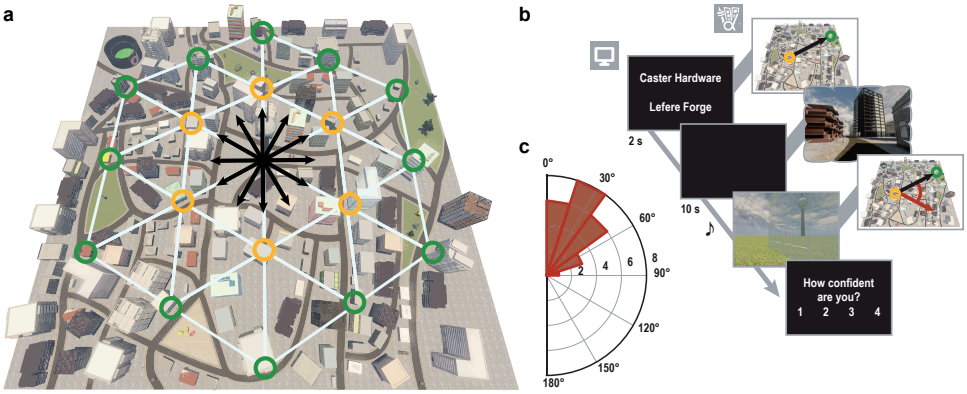


Figure 1 | Direction-imagination task.

(**a**) Twelve evenly spaced directions were sampled using 18 buildings distributed regularly across Donderstown. We sampled each direction (indicated by black arrows) from different start locations (yellow circles), which dissociated the directions from visual features of imagined views (Figure 1 - figure supplement 2), and employed a counterbalancing regime ensuring equal sampling of directions and start locations throughout the experiment (see Materials and methods). Buildings marked with a green circle served as target locations only. Importantly, the regular arrangement of building locations did not correspond to the street layout and was not revealed to participants, who experienced Donderstown only from a first-person perspective (see also Figure 1 - figure supplement 1d). (**b**) Trials began with a cue indicating start (top building name) and target (bottom building name) location and thereby defining the relevant direction (black arrow). During an imagination period the screen was black and participants were instructed to imagine the view they would encounter when standing in front of the start building facing the direction of the target building. An auditory signal terminated the imagination period and participants indicated the imagined direction (red arrow) in a sparse VR environment, followed by a confidence judgment. Performance was measured as the absolute angular difference between the correct and the indicated direction (red arc). Note that only the bottom row of images was presented to participants, top row for illustration only. (**c**) Circular histogram of average absolute angular difference between correct and indicated directions across participants (mean error 33.68° ± 19.09° SD).

RESULTS

We combined fMRI with multi-voxel pattern analysis and VR to investigate whether the entorhinal grid system contributes to the imagination of directions from stationary viewpoints (Figure 1a and Figure 1b). After extensive navigation training (see Materials and methods and Figure 1 - figure supplement 1), participants were asked to imagine directions between pairs of buildings in 'Donderstown', a large-scale realistic VR city (www.doellerlab.com/donderstown/). In a carefully counterbalanced design, we probed the fine-grained representations of twelve equally spaced directions (see Materials and methods and Figure 1 - figure supplement 2). Imagined directions had to be indicated (Figure 1b) and participants successfully performed this task (mean error 33.68° ± 19.09° SD; Figure 1c). Behavioral performance in the direction-imagination task was highly correlated with navigation success ($r = 0.85$, $p < 0.001$) and the accuracy of direction estimates ($r = 0.94$, $p < 0.001$) during training and performance in a post-scan map test ($r = 0.95$, $p < 0.001$), indicating successful translation of the acquired representation of Donderstown to the imagination task (Figure 1 - figure supplement 3).

The contribution of spatial representations to imagination was assessed using representational similarity analysis (Kriegeskorte and Kievit, 2013), which compares activation patterns across voxels to estimate neural similarity. In line with the suggested role for grid cell computations in vector navigation (Bush et al., 2015), we expected the grid-cell system to be involved in determining the vector comprising the direction and distance from the start to the target building in our task. Our approach focused on the direction to the target to track putative grid-cell representations during imagination by systematically comparing neural similarity of imagined directions with varying angular differences (see Materials and methods). We predicted that an involvement of the grid system in mental simulation should be reflected in a 60°-periodic pattern-similarity structure in the entorhinal cortex, consistent with the hexagonal firing properties of grid cells (Hafting et al., 2005) and the hexadirectional fMRI signal in entorhinal cortex observed during virtual navigation (Doeller et al., 2010; Horner et al., 2016; Kunz et al., 2015). It is important to note that we did not rely on the building layout as an absolute reference frame in our analyses, but rather assessed pattern similarity based on the *relative* angle between the directions sampled in a trial pair, see below.

In a first step, we ascertained that absolute directional representations are detectable with our novel imagination task. We expected increased neural similarity

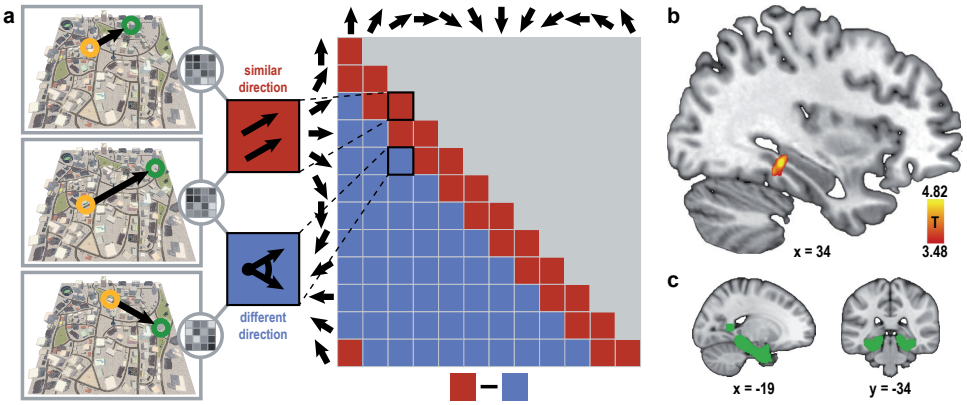


Figure 2 | Absolute directional coding in posterior parahippocampal gyrus.

(**a**) Analysis logic of the *one-fold* directional analysis for three example trials. High pattern similarity was predicted for pairs of trials sampling similar directions with a maximum angular difference of 30° (red) compared to trials sampling directions 60° or more apart (blue). Note that for illustration purposes the predicted similarity matrix is shown for comparisons across conditions, not single trials. (**b**) Searchlight results show a significant cluster of voxels in the posterior parahippocampal gyrus (peak voxel MNI coordinates: 34 -34 -10; T_{23} = 4.82, p = 0.024 corrected for multiple comparisons using small-volume correction) with higher pattern similarity for trials sampling similar directions compared to trials sampling dissimilar directions. Results are shown on the structural MNI template. For display purposes, the statistical map is thresholded at p < 0.001 uncorrected. (**c**) Sagittal and coronal view of the mask used to correct for multiple comparisons (see Materials and methods) displayed on the MNI template brain.

during imagination of similar directions, consistent with previous findings of absolute directional coding during navigation in parahippocampal cortex (Doeller et al., 2010) and two recent studies reporting directional representations during imagination in a local reference frame in the retrosplenial complex (Marchette et al., 2014) and coarse representations of directions to a goal in the entorhinal/subicular region (Chadwick et al., 2015). However, it remains unclear whether global spatial representations are involved in human imagination in the absence of visual input.

Here, we compared pattern similarity during imagination of directions in pairs of trials sampling similar directions (angular difference ≤ 30°) to pairs of trials sampling dissimilar directions (Figure 2a) in brain regions representing facing direction (Baumann and Mattingley, 2010; Chadwick et al., 2015; Marchette et al., 2014; Vass and Epstein, 2013, 2016). We observed the predicted one-fold symmetric pattern-similarity structure in a cluster of voxels in the left posterior parahippocampal gyrus (T_{23} = 4.82, p = 0.024, FWE-corrected for multiple comparisons using small volume correction; Figure 2bc; see Materials and methods). Increased pattern similarity for similar directions was not due to trial comparisons with identical building combinations (Figure 2 - figure supplement 1). Further, this effect was not

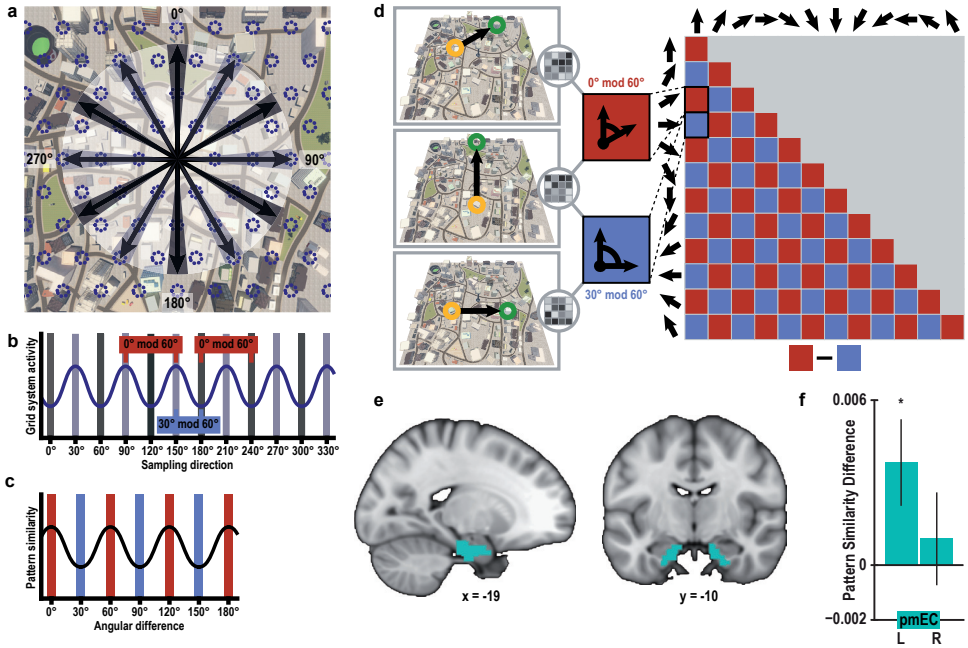


Figure 3 I Grid-like representations during imagination.

(a) Six-fold symmetric firing fields of a hypothetical grid cell (dark blue dotted circles) superimposed on an aerial view of Donderstown. Black arrows indicate the twelve sampled directions; light and dark shading highlights directions (multiples of) 60° apart. For illustration purposes, the grid orientation is aligned to the sampled directions; see Figure 3 - figure supplement 1 for a different example. **(b)** The firing rate of the hypothetical response of the grid-cell system as a function of direction, showing a 60° modulation. Shading displays sampling of directions and red and blue markers indicate the two conditions. Note that the oscillatory firing pattern is sampled at the same phase in the 0° modulo 60° condition, but at different phases in the 30° modulo 60° condition. **(c)** Based on this, we expected a 60° modulation of fMRI pattern similarity values when comparing trial pairs based on the angular *difference* of their sampled directions. Red and blue shading illustrates the two conditions. **(d)** Specifically, we predicted higher pattern similarity for trial pairs with a remainder of 0° (0° modulo 60° condition, red) compared to trial pairs with a remainder of 30° (30° modulo 60° condition, blue), when dividing the angular difference of the pair's sampling directions by 60°. Note that for illustration purposes the predicted similarity matrix is shown for comparisons across conditions, not single trials. **(e)** ROI mask for posterior medial entorhinal cortex (pmEC) from previous report (Navarro Schröder et al., 2015). **(f)** Pattern similarity difference (mean and S.E.M.) between the two conditions. The left pmEC exhibited a significant 60° modulation of pattern similarity. No significant differences in pattern similarity were observed in the right pmEC (T_{23} = 0.57, p = 0.58).

driven by the specific locations used to sample directions or the distances between these locations in Donderstown (Figure 2 - figure supplement 2, see Materials and methods).

Having verified that we can detect directional representations in our novel imagination paradigm, we tested, in a next step, whether activation patterns during imagination follow a six-fold rotational symmetry, akin to the six-fold symmetric firing

pattern of grid cells (Hafting et al., 2005) and the six-fold modulation of entorhinal fMRI signals during virtual (Doeller et al., 2010; Kunz et al., 2015) and imagined (Horner et al., 2016) navigation in humans. The rationale underlying our analysis is that activation patterns during directional imagination should exhibit the highest neural similarity for directions that are (multiples of) 60° apart from each other (see Figure 3a-d and Figure 3 - figure supplement 1 for details of analysis logic). Because grid cells are most abundant in the medial entorhinal cortex in rodents (Hafting et al., 2005), we predicted the effect to be present in posterior-medial entorhinal cortex (pmEC), the likely homologue region of the rodent medial entorhinal cortex in the human brain (Navarro Schröder et al., 2015) (Figure 3e).

We observed pattern similarity increases with a 60° periodicity in the left pmEC (T_{23} = 2.37, p = 0.027; one-tailed test, Bonferroni corrected for test in both hemispheres; Cohen's d = 0.48; Figure 3f and Figure 3 - figure supplement 2b). The effect was further confirmed using permutation-based significance testing (pseudo T23= 2.89, p = 0.008; see Materials and methods). A control analysis showed that the effect was not present in the anterior lateral entorhinal cortex (p > 0.9; Figure 3 - figure supplement 2; see Figure 3 - figure supplement 3 for information on signal quality in the entorhinal cortex), the human homologue of lateral EC, which does not contain grid cells (Hafting et al., 2005). The 60° periodicity in left pmEC was consistent across all angular differences (Figure 3 - figure supplement 4 and Figure 3 - figure supplement 5) and the effect was not driven by the specifics of our design and the VR town used. Specifically, the effect remained significant after excluding combinations of trial pairs (Figure 3 - figure supplement 6) with the same start (T_{23} = 2.39, p = 0.025) or target location (T_{23} = 2.57, p = 0.017), the same combination of start and target location (T_{23} = 2.45, p = 0.022) and comparisons from the same task block (T_{23} = 2.08, p = 0.049). Further control analyses demonstrated that the effect was independent of the mean distance between start and target locations in a trial pair (T_{23} = 2.37, p = 0.027; Figure 3 - figure supplement 7), the difference of this distance within a pair (T_{23} = 4.32, p < 0.001) and the mean distance between all four buildings in a given trial pair (T_{23} = 2.37, p = 0.027). Behavioral performance did not differ between the conditions (T_{23} = 1.24, p = 0.227, Figure 3 - figure supplement 8).

Furthermore, the effect was specific to a 60° modulation of pattern similarity values and there was no evidence for coding of cardinal directions in the entorhinal cortex (Figure 3 - figure supplement 9; see also Materials and methods). Results of a whole-brain searchlight analysis confirmed the 60° periodicity of pattern similarity increases in pmEC observed in the ROI analysis (T_{23} = 4.04, p = 0.046, FWE-corrected for

multiple comparisons using small volume correction; Figure 3 - figure supplement 10). A similar pattern similarity structure was observed in regions in parietal and visual cortices (Figure 3 - figure supplement 10), which might reflect reactivation of egocentric and visual representations associated with imagined directions, possibly modulated by entorhinal representations in line with a model of mental imagery (Bird et al., 2012; Byrne et al., 2007). Future research will need to investigate these putative interactions in more detail.

DISCUSSION

In sum, we report two important findings: Firstly, pattern similarity values in the parahippocampal gyrus exhibited a one-fold symmetry congruent with fine-grained representations of imagined facing direction, reflecting the role of this brain region - which has been implicated in spatial processing in the absence of visual input (Wolbers et al., 2011) - in representing the directional aspect of the imagined views. An alternative explanation of this effect through visual similarity of the imagined views appears unlikely due to the complex nature of the task in which each direction was sampled from multiple locations in our large-scale environment (Figure 1 - figure supplement 2) and buildings served as cues for a wider range of sampling directions. Therefore, our finding provides the first evidence for fine-grained coding of absolute direction at an unprecedented angular resolution of 30°, consistent with the characteristics of the head direction system in rodents (Taube et al., 1990a; Taube, 2007), and constitutes a three-fold increase in resolution of the directional representations observed in humans compared to previous studies (Marchette et al., 2014; Chadwick et al., 2015; Baumann and Mattingley, 2010; Vass and Epstein, 2013, 2016; Shine et al., 2016). Secondly, the structure of pattern similarity in entorhinal cortex was characterized by a six-fold rotational symmetry akin to the firing properties of grid cells (Hafting et al., 2005). Our findings provide evidence for an involvement of grid-like representations in mental simulation in the absence of imagined movement.

Crucially, participants imagined directions from stationary viewpoints in a realistic, large-scale virtual city and were not re-exposed to the virtual town during the imagination task. Therefore, our findings provide novel evidence, complementary to a recent report (Horner et al., 2016) showing evidence for grid-like entorhinal processing during imagined movement through a simple virtual arena. Furthermore, in contrast, we investigated spatial processing in a large-scale, urban environment

(Stokes et al., 2015) and, moreover focused on multi-voxel patterns. In particular, we demonstrate that this novel analysis approach, which does not rely on the estimation of the orientation of the hexadirectional signal in entorhinal cortex in an independent data set (Constantinescu et al., 2016; Doeller et al., 2010; Horner et al., 2016; Kunz et al., 2015), is sensitive to grid-like entorhinal signals by capitalizing on the six-fold symmetry of grid cell firing patterns. Contrary to the previously employed approach relying on the estimation of the orientation of the hexadirectional signal for each participant (Constantinescu et al., 2016; Doeller et al., 2010; Horner et al., 2016; Kunz et al., 2015), the individual grid orientation is not approximated using the multivariate analysis. Yet, the grid orientation might influence the strength of the grid-like entorhinal signal observed in a given participant because the sampled directions might be more or less aligned with this individual's grid orientation (Figure 3a-c and Figure 3 - figure supplement 1 for illustration). This needs to be taken into consideration when aiming to relate grid-like signals to behavior. However, only the multivariate approach enabled us to investigate the six-fold rotational symmetry in our large-scale environment, in which a continuous sampling of directions as required for the estimation of the orientation of the hexadirectional signal would not have been feasible. This parsimonious approach might prove valuable for future studies investigating the role of grid-like signals in human cognition, in particular in studies with children (Bullens et al., 2010) or older participants (Schuck et al., 2015a) and in clinical settings (Hartley et al., 2007; Maguire et al., 2001), where time for data acquisition is typically limited and could for instance help to further understand the putative link between the entorhinal grid system and Alzheimer's disease (Kunz et al., 2015).

On a theoretical level, our findings are consistent with accounts of imagination, positing medial-temporal-lobe involvement in the reactivation and recombination of prior experiences (Buckner, 2010; Byrne et al., 2007; Hassabis and Maguire, 2007; Hassabis et al., 2007a; Hasselmo, 2009; Schacter et al., 2012). The hippocampal formation and grid cells in particular have been implicated in path integration (Hafting et al., 2005; Wolbers et al., 2007), a central aspect of which is computing a homing vector based on translations from a given starting point (Vickerstaff and Cheung, 2010). Notably, the grid system is well-suited to also perform the inverse operation of calculating relative vectors between known positions in the service of navigational planning (Bush et al., 2015). Hence, it is plausible that the grid-cell system contributes to the calculation of vectors between start and target location during imagination (Bird et al., 2012; Bush et al., 2015; Hasselmo, 2009; Horner et al., 2016), while the head direction system (Taube, 2007; Taube et al., 1990a)

processes the absolute direction between the two locations (Bird et al., 2012; Byrne et al., 2007; Hasselmo, 2009) in our task. Our findings suggest an involvement of the entorhinal grid system in calculating vectors to target locations during navigational planning, in line with a theoretical account of vector navigation (Bush et al., 2015).

Functional neuroimaging can measure the firing pattern of specific cell types only indirectly (Logothetis, 2008). However, intracranial recordings in patients exploring virtual-reality environments demonstrated the existence of place (Ekstrom et al., 2003; Jacobs et al., 2010) and grid (Jacobs et al., 2013) cells in the human hippocampus and entorhinal cortex, respectively. Importantly, our results are in line with single-cell recordings in rodents that suggest a possible contribution of spatially tuned cells to future anticipation via place cell preplay of upcoming trajectories (Dragoi and Tonegawa, 2011) and preferential preplay of firing sequences of paths leading to motivationally relevant locations (Ólafsdóttir et al., 2015). Prospective coding properties of grid cells (De Almeida et al., 2012; Kropff et al., 2015) and recent evidence for grid cell replay (Ólafsdóttir et al., 2016) further suggest a similar involvement of the entorhinal grid system in future anticipation and prediction. By translating these ideas to human imagination, during which content-specific firing rate increases of neurons in the human medial temporal lobe have been observed (Kreiman et al., 2000), it is conceivable that spatially tuned cells provide the machinery for the flexible recombination of spatial and mnemonic details necessary for the construction of mental simulations (Bird et al., 2012; Brown et al., 2016; Buckner, 2010; Byrne et al., 2007; Eichenbaum and Cohen, 2014; Hassabis and Maguire, 2007; Hassabis et al., 2007a; Hasselmo, 2009; Schacter et al., 2012) and the representation of conceptual knowledge (Constantinescu et al., 2016).

In concert with the recent report of grid-like processing in the entorhinal cortex during imagined navigation (Horner et al., 2016) our findings provide a substantial advancement for the field. Importantly, grid-like entorhinal signals during imagined navigation were observed with a similar orientation as during actual navigation through the VR environment (Horner et al., 2016). This finding strengthens our interpretation of the six-fold symmetric pattern similarity structure in the entorhinal cortex during imagination of directions from stationary viewpoints observed in this study as reflecting computations of the entorhinal grid system operating similarly in our realistic large-scale VR city as during navigation in smaller and simpler environments typically used in electrophysiological recording studies in rodents (Hafting et al., 2005) or fMRI experiments in humans (Doeller et al., 2010; Horner et al., 2016; Kunz et al., 2015). Importantly, the interpretation of our results as a

global grid signal coding for space beyond boundaries and obstacles is in line with the report of a global grid pattern emerging with experience in rodents exploring an environment divided into two connected compartments (Carpenter et al., 2015).

In conclusion, we show involvement of both absolute directional parahippocampal and grid-like entorhinal signals in imagination, which provides important evidence for these representations in the absence of sensory input or imagined movement. This might suggest a more fundamental role of spatial computations in the grid-cell system during mental simulation and possibly other forms of prospective coding and future thinking in the service of goal-directed decision-making (Bird et al., 2012; Buckner, 2010; Byrne et al., 2007).

MATERIALS AND METHODS

Participants

32 male participants were recruited via the online recruitment system of Radboud University Nijmegen. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, The Netherlands) and participants gave their written informed consent prior to the experiment. All participants had normal or corrected to normal vision and were compensated for their participation. Eight participants were excluded from the analysis because of motion sickness during the VR navigation training (n = 2), technical problems with the MRI scanner (n = 1) or chance-level performance during the direction imagination task (n = 5; median absolute angular error not significantly smaller than 90° as determined by Wilcoxon signed-rank test). Thus, 24 participants (age range 18-29 years, mean age 24.52 years, standard deviation 2.91 years) entered the analysis.

Procedure

The experiment was conducted on two days and consisted of an extensive behavioral training in our virtual-reality city 'Donderstown' (http://www.doellerlab.com/donderstown/) and a direction imagination task in the MRI scanner. Additionally, participants' ability to locate the task-relevant buildings on a map was assessed. On the first day, participants learned the names and locations of 18 buildings in Donderstown in a two-hour training session in the virtual city. Before the fMRI session on the following day, participants were trained for an additional hour.

VR city

Donderstown provides a large-scale, realistic urban environment comprising a complex layout of streets, squares and parks. Inspired by the street-map of a medievally founded German town, the layout of Donderstown resembles the irregular outline of a typical European city with curved roads. The streets of the city were not named. Participants were instructed to use the centrally positioned radio tower for orientation along with the mountain range surrounding the city. The skybox was rendered at infinity and the sun as the major cue of directional lighting was placed at the zenith above the center of the city. Thus, shadows could not serve as directional cues during any part of the experiment. 3D models of buildings for the city were created using 3ds Max (version 2014, www.autodesk.com/products/3ds-max/). The city was built and presented using the Unreal Development Kit (v2013-07, http://udn.epicgames.com/Three/WebHome.html).

Crucially, the task-relevant buildings were placed at specific locations in Donderstown allowing us to sample twelve evenly spaced directions in the imagination task (Figure 1a) and to dissociate the start location from the sampled direction. Specifically, the entrance doors of these 18 buildings were approximately located at the vertices of equilateral triangles, which were arranged in a hexagon. Note that participants never saw a top-down view of the city, but only experienced the city from a first-person view and that the layout of streets and landmarks did not correspond to the regular pattern of buildings relevant for the direction imagination task. Therefore it is unlikely that participants were aware of the regular layout of task-relevant buildings (see Figure 1 - figure supplement 1d).

Behavioral Training

Familiarization phase. On the first day of training, participants were familiarized with Donderstown and the controls of the computer game by freely navigating the city for ten minutes in search for a set of landmarks spread across the city (see Figure 1 - figure supplement 1a). Participants first encountered Donderstown at a position close to the city's center facing the radio tower located in the center of the city. This position and orientation is indicated in Figure 1 - figure supplement 1a and also served as the start location during the delivery task described below. For the familiarization phase, seven landmarks (e.g. a statue) distributed across the city were selected and participants were instructed to look for these landmarks while exploring the city. Donderstown was presented on a computer screen at a resolution of 1024 x 768 pixels. Participants experienced Donderstown exclusively from a first-person perspective and navigated using the mouse and keyboard to control the player movements.

Association task. Afterwards, participants learned the names of the 18 task-relevant buildings, with names randomly assigned to the buildings for each participant. We used names of unknown US-based companies consisting of two or three words. Participants were asked to remember the building-name associations during learning blocks, in which the image of a building was presented together with its name using Presentation (version 16.4, www.neurobs.com/presentation; Figure 1 - figure supplement 1b). Learning blocks were followed by blocks of test trials during which participants had to select the building corresponding to a presented name from a set of three building images within 4 s. After each test block, participants received feedback about their performance in that block (percentage of correct answers). After reaching a learning criterion of three error-free test blocks (mean 7.00 ± 3.27 SD blocks to criterion), participants were trained in Donderstown for the rest of the 2-h training session as described below.

Delivery task. In order to learn the building locations in the city, participants completed a delivery task (average length of navigation training 80.32 min ± 8.00 SD min), in which the goal was to find the target building whose name was presented on the screen (Figure 1 - figure supplement 1c). To familiarize participants with the building locations, participants were guided to the buildings by following pylons until having visited each building three times. After this initial phase, participants searched for the buildings themselves, which were targeted in random order. If a participant could not find a building during the first 36 trials of this phase, a help button made a widely visible red arrow appear in the sky above the target building. We determined navigation performance by the number of buildings found after the end of the guided navigation phase. Upon finding a building, the player position was fixed for 7 s to allow time for the encoding of the building's location. Navigation was restricted to rotations to look around during this time period.

Direction training. To emphasize the relevance of the directional relationships between buildings, participants were prompted to estimate the direction to the next target building after having found a building (Figure 1 - figure supplement 1c). The name of the new target building was presented and participants were instructed to rotate until facing the new target. Performance on this task was measured as the absolute angular difference between the correct and the indicated direction. No explicit feedback on the direction was given, but participants were instructed to verify their direction estimates by searching for the target building in the indicated direction.

During the second training session, which took part approximately 24 hours after the first, participants completed a shorter version (1 h) of the training procedure described above. First, participants were again trained on the building names using the *Association task* described above. The learning criterion was reduced to one error-free test block (mean 1.23 ± 0.47 SD blocks to criterion). Again, participants spent the remainder of the training session performing the *Delivery task* (average length of navigation training 40.44 min ± 5.09 SD min). The setup for the navigation training on the second day was identical to the previous day with the only difference that participants were guided to each building by pylons only once. Again, we measured navigational success as the number of buildings found after the guided navigation period as well as the accuracy of the direction judgments obtained.

To test the relationship of performance during training to behavioral performance in the MRI direction imagination task, the number of buildings found was aggregated across both training sessions and the accuracy of direction judgments was averaged across all estimates. The close relationship between training performance and direction estimates in the imagination task (Figure 1 - figure supplement 3) was not due to the minor differences in the length of the navigation training between participants. Following the second training session, participants were instructed about the direction imagination task for the subsequent fMRI session and performed three practice trials in order to ensure all aspects of the task were understood correctly. All training sessions were conducted in a behavioral laboratory.

fMRI direction imagination task

While undergoing fMRI, participants performed a novel direction imagination task. The task consisted of four blocks of 24 trials. In each trial, participants were cued with the names of two buildings and were instructed to close their eyes and imagine themselves standing in front of the start building facing the direction of the target building. To successfully imagine views from variable locations defined by the relative positions of buildings without re-exposure to Donderstown participants had to make use of their allocentric representation of the city. The angle of the vector between start and target building defined the correct direction in each trial with respect to the coordinates of Donderstown and defined the trial conditions for the representational similarity analysis (RSA) (Kriegeskorte and Kievit, 2013). The two building names were presented simultaneously for two seconds.

The layout of the buildings relevant for this task enabled the controlled sampling of twelve directions with an angular difference of 30° between adjacent directions

(Figure 1a). Using the inner six buildings as start buildings for the direction imagination task allowed for the dissociation of start location and sampled direction. From each of the six start buildings ten of the twelve directions were sampled (Figure 1 - figure supplement 2). Since the 60 unique building combinations oversampled the directions aligned with the main axes of the hexagon of buildings, we randomly excluded two trials sampling each of these directions with the constraint to exclude two building combinations per start location. The remaining 48 unique building combinations sampled each direction four times and were used for two task blocks of 24 trials. For the two blocks, trials were randomly drawn from the available combinations of start and target building with the constraints that all directions were sampled twice and each start building served as a start location four times in a block. The order of trials was randomized within blocks. For the second set of two blocks we followed the same randomization procedure.

Participants were instructed to imagine the direction between the two buildings with their eyes closed for 10 s during which the screen was black. Then they were prompted by a tone to open their eyes and indicate the imagined direction. For this behavioral response, a sparse VR environment was used (Figure 1b). The city was replaced by a grass plain with the only remaining cues for orientation being the radio tower, which had been located in the center of Donderstown during navigation and the mountain range surrounding the city. One of the mountains on the south-east side of the city had a summit cross on it, which had also been visible during training. In each trial, the start building was also presented on the grass plain at the position corresponding to the building's location in Donderstown. The participant was spawned at the imagined location in front of the start building with the task to indicate the previously imagined direction to the target building by rotating until a red cross presented centrally on the screen was pointing in this direction. In order to facilitate the estimation of directions to targets behind the start building, the building was turned transparent after 2 s in every trial. The accuracy of the behavioral response was measured as the angular distance between the correct direction defined by the start and target building and the response direction (Figure 1bc). The magnitude of errors observed in the behavioral responses was consistent with performance levels reported in previous studies using similar tasks (Schinazi et al., 2013; Zhang et al., 2014).

Each trial ended with a confidence rating on a four-point scale and was followed by an inter-trial-interval (randomized length of 1.8, 3.6 or 5.4 s). Three trials within a block served as catch trials in which a snapshot from the VR city was presented after

the imagination period, but before the onset of the response screen. This snapshot was taken from the imagined location in front of the start building and participants had to indicate within 2 s after the end of the imagination phase whether it was facing the direction of the target building or a direction different by 90°. In 50% of the catch trials, the snapshot showed the correct view. We introduced these trials in order to emphasize the relevance of vivid imagination of the relevant views. Participants responded significantly above chance for the catch trials in which they responded in time (T_{23} = 2.70, p = 0.013).

Map test

After completion of the direction imagination task, participants' memory for the building locations was tested. In each trial, the name of a building was presented together with a map showing the outline of the city's streets. The objective of the task was to move a cursor to the location where the building was thought to be positioned in the city. In order to facilitate orientation on the map, the landmarks participants searched during the familiarization phase were represented by symbols on the map. Participants were instructed about the identity of these symbols. Every building was to be located once. Participants' accuracy was defined as the Euclidean distance between the correct building location on the map and the location indicated by the participant.

MRI data acquisition

Functional images were acquired at a 3 T Siemens Trio MRI system (Siemens, Erlangen, Germany) using a 3D EPI sequence with an isotropic voxel size of 2mm and a TR of 1800 ms (TE = 25 ms, 64 slices, distance factor 50%, flip angle 15°, field of view 224x224x128 mm). Four runs corresponding to the four task blocks were collected. The duration of each run depended on the time required by the participant to complete the task block. On average, a run lasted 11.76 min (± 1.63 SD min). T1-weighted structural images were acquired with a MPRAGE sequence (TR = 2300 ms, TE = 3.03 ms, flip angle 8°, in-plane resolution = 256x256 mm, voxel size 1 mm isotropic).

MRI data analysis

Preprocessing

Preprocessing of the functional images of each of the four runs was carried out using the FSL toolbox (version 5.0.4, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). The images were high-pass filtered (cut-off: 100 s) and motion-corrected. For the RSA

each run was linearly registered to the participant's structural scan, which served as a common reference space for functional scans from all four blocks and was downsampled to an isotropic voxel size of 2 mm to correspond to the voxel size of the functional scans. A gray matter segmentation was carried out to obtain a gray matter mask for use in the analysis described below.

Representational similarity analysis: absolute directional coding

RSA (Kriegeskorte and Kievit, 2013) was implemented using custom Matlab (version R2014a, www.mathworks.com/matlab/) scripts. For the preprocessed time series of every voxel, a general linear model (GLM) was calculated with the motion parameters obtained during preprocessing as predictors. The residuals of these GLMs (i.e. what could not be explained by the motion parameters) were then used for representational similarity analysis.

For this analysis, three volumes corresponding to the peak of the BOLD response of the second half of the imagination time window were averaged for each trial. Specifically, the analysis time window consisted of the volume during which the 10 s imagination period ended and the two following volumes. This time period was chosen due to the complex nature of the task, which required the participant to first retrieve the respective buildings and their locations before being able to imagine the direction between them. To rule out a potential influence of brain activity corresponding to the onset of the response screen on the analyzed volumes we conducted a whole-brain searchlight analysis (with the parameters described in detail below), which compared trial pairs with identical visual input at the onset of the response screen to pairs where this visual input differed. Since the initial response screen always showed the start building, trial pairs with the same start location would exhibit increased pattern similarity in brain areas processing visual information if the volumes analyzed contained brain activity corresponding to visual processing. No clusters showed this effect at a significance threshold of $p < 0.001$ uncorrected and a cluster extent threshold of 10 voxels.

We performed the searchlight analyses using search spheres with a diameter of 7 voxels (1.4 mm). This approach iteratively calculates pattern similarity values for search spheres centered on the current voxel. All within brain voxels could be the center of a sphere, but within each sphere, the analysis was restricted to gray matter voxels by using the mask obtained from the structural image segmentation during preprocessing. Only search spheres including more than 30 voxels were analyzed. Pattern similarity was calculated as the Pearson correlation coefficient across all

voxels within a sphere, separately for all possible pairs of trials. To test the hypothesis of increased representational similarity for similar directions, we averaged Fisher z-transformed correlation coefficients for two different conditions: Pairs of trials probing the same direction or directions with an angular difference of 30° were contrasted against pairs of trials sampling directions 60° or more apart (Figure 2a). We chose to investigate representations of absolute directions with a resolution of 30° based on the firing properties of head direction cells (Taube, 2007), which might underlie an absolute directional code in our task, and to achieve sufficient statistical power. The difference in mean pattern similarity between these conditions was calculated for each search sphere on the single subject level.

The resulting difference images were normalized (non-linear registration) to the MNI template (2 mm resolution) using FSL and smoothed using a Gaussian kernel of 4mm full width at half maximum. On the group level, significance was assessed by performing a one-sample T-test in SPM 8 (http://www.fil.ion.ucl.ac.uk/spm/) on the difference images obtained at the subject level (cluster extent threshold 10 voxels). The resulting statistical map is displayed at 0.5 mm resolution on the MNI template in Figure 2b. Small volume correction was performed based on our a priori hypothesis about the brain regions involved in the representation of facing direction in humans as reported in previous studies (Baumann and Mattingley, 2010; Chadwick et al., 2015; Doeller et al., 2010; Marchette et al., 2014; Vass and Epstein, 2013). Therefore, a mask (Figure 2c) containing the retrosplenial complex, parahippocampal gyrus, entorhinal cortex and the subiculum was constructed from thresholded masks of these regions. For the subiculum and the parahippocampal gyrus masks were selected from the Jülich Histological atlas (Eickhoff et al., 2007) and the Harvard-Oxford Cortical Structural atlas (Desikan et al., 2006) distributed with FSL, respectively. The retrosplenial complex was defined based on peak voxel coordinates from a previous study in which participants imagined directions in a local reference frame (Marchette et al., 2014). We created a spherical mask with a radius of 6 mm around the peak voxel in the right retrosplenial complex (MNI coordinates: 14 -58 11) and added a second sphere with the same radius in the same location in the left hemisphere. Masks for the posterior medial and anterior lateral parts of the entorhinal cortex were based on a previous study (Navarro Schröder et al., 2015). The mask used for small volume correction is displayed at 1 mm resolution in Figure 2c.

Absolute directional coding: Control analyses

Excluding trial pairs with the same start and target location. The parahippocampus is known to be involved in the processing of visual scenes (Epstein, 2008). Therefore, increased pattern similarity for similar directions could potentially be driven by trial pairs in which participants were cued to imagine the identical spatial scene. We performed a searchlight analysis as described above and excluded trial pairs from the analysis, which sampled a direction with the identical combination of buildings. On the group level, we observed an effect in a very similar location in the left posterior parahippocampal gyrus (Figure 2 - figure supplement 1).

Exclusion of trial pairs with the same start location. To control for the possibility that increased pattern similarity was due to pairs of trials sampling directions from the same start location, we excluded these comparisons from the analysis. Specifically, we extracted the matrix containing all pair-wise correlation coefficients from the sphere centered on the peak voxel of the cluster observed in our main analysis (Figure 2). To this end we registered the peak voxel coordinate to each subject's structural image (2 mm resolution), which again served as the common space for the analysis of the functional images, and computed the pair-wise correlations for all possible pairs of trials in the sphere centered on this voxel (same parameters as for main searchlight analysis). After excluding all comparisons of trial pairs sampling directions from the same start location, we computed the difference in mean pattern similarity between pairs sampling similar (angular differences ≤ 30°) and dissimilar (angular differences ≥ 60°) directions as in our main analysis. We still observed increased pattern similarity for pairs of trials sampling similar compared to dissimilar directions (T_{23} = 3.01, p = 0.006; Figure 2 - figure supplement 2a, bar I).

Exclusion of trial pairs with the same target location. Following the same approach, we also excluded the possibility that increased pattern similarity for similar directions was driven by more frequent imagination of the same target buildings. As shown in Figure 1 - supplement 2, directions were sampled from the inner ring of buildings in Donderstown and mostly targeted locations towards the outside of the city. Therefore, target buildings could not be sampled from all directions. After excluding comparisons of trials targeting the same location in a second analysis, pattern similarity remained greater for comparisons of similar directions (T_{23} = 4.07, p < 0.001; Figure 2 - figure supplement 2a, bar II).

Rationale for controlling distance-related effects. In every trial, the direction to be imagined was defined as the angle of the vector between the start and target building.

In this set of control analyses we additionally considered the Euclidean distance between the start and the target location. We adopted three different approaches to combine the distance information for a given trial pair and to obtain a distance matrix with the distance for each comparison (Figure 2 - figure supplement 2b). To test whether distances differed between conditions we computed the same contrast as in our main analysis: For each subject we calculated the difference between the mean distance of trial pairs sampling similar directions and the mean for pairs sampling dissimilar directions and tested these differences against 0 using a one-sample t-test. The results are described below. As described above, we obtained the pair-wise similarity matrix from the sphere centered on the peak voxel of the cluster from the main analysis (Figure 2). In each analysis, we then computed a GLM with pairwise distances as a continuous predictor and the pairwise Fisher z-transformed correlation coefficients as the criterion for each participant. The residuals of these GLMs (i.e. pattern similarity that could not be explained by the distance predictor) were then used to compute the pattern similarity differences between trial pairs sampling directions with a maximum angular difference of 30° and pairs sampling directions with an angular difference of 60° or more. After all analyses, pattern similarity remained significantly greater for trial pairs sampling similar directions compared to trial pairs sampling dissimilar directions. The details of these analyses are described in the following.

Controlling for mean distance of trial pairs. First, we averaged the Euclidean distances between start and target building in a trial pair. The mean Euclidean distance was greater for pairs sampling directions with a maximum angular difference of 30° (T_{23} = 8.30, $p < 0.001$). By accounting for the mean length of the two vectors, we excluded a potential influence of imagined distance to the target location. Pattern similarity remained increased for trial pairs sampling similar directions (T_{23} = 3.90, $p < 0.001$, Figure 2 - figure supplement 2c, bar I).

Controlling for different distances within trial pairs. Second, we considered a modulation of pattern similarity based on the difference in distance between start and target location of the trials defining a pair. Trial pairs with angular differences of maximally 30° had more different distances (T_{23} = 58.23, $p < 0.001$). After controlling for the difference in length of the two trial vectors in a pair, pattern similarity was still increased for pairs of trials sampling similar directions (T_{23} = 3.61, $p < 0.001$, Figure 2 - figure supplement 2c, bar II).

Controlling for the 'neighborhood distance'. Lastly, we quantified a 'neighborhood' distance, which we defined as the mean of the Euclidean distances of the six vectors connecting all four buildings in a trial pair. This distance was low when both start and both target buildings were in the same part of Donderstown. These distances were smaller for pairs sampling similar directions (T_{23} = -77.51, $p < 0.001$). After controlling for the fact that buildings used to sample similar directions were located more closely together in Donderstown, pattern similarity remained increased for pairs of trials sampling similar directions (T_{23} = 3.97, $p < 0.001$, Figure 2 - figure supplement 2c, bar III). This rules out the possibility that sampling in similar parts of Donderstown is responsible for increased pattern similarity for trial pairs sampling similar directions.

Visual similarity of imagined views. Based on the well-established role of the parahippocampal cortex in scene processing, increased pattern similarity for similar directions in this region could potentially be driven by increased visual similarity of the scenes imagined by the participants. Even though this appears unlikely due to the sampling of each direction from multiple locations (Figure 1 - figure supplement 2) we compared the visual similarity of the scenes to be imagined based on a set of visual features (Milivojevic et al., 2015). For each trial, the view from the start position facing the direction of the target building was captured at a resolution of 1024x768 pixels. Note that participants did not see these images during the task, but were instructed to imagine the views. Five statistics were used to quantify visual features of each image: By computing a two-dimensional discrete Fast-Fourier-Transform we obtained magnitude and phase metrics of the images. Next, the red, green and blue values for each image were converted to CIELAB color space. This color space provides information about the luminance and two color opponent dimensions (red-green and yellow-blue) corresponding to the cone responses of the human retina. Each of the five image statistics was vectorized and Pearson correlations were computed to quantify visual similarity across pairs of images.

We compared visual similarity for pairs of trials sampling similar directions (treating angular differences of 30° or less as similar akin to the absolute directional coding analysis) from different start positions. For each participant, we performed a two-sample t-test on the Fisher z-transformed correlation coefficients of each image dimension. While we did not observe significant differences for the magnitude information or the color dimensions (all Bonferroni-adjusted $p > 0.2$), phase and luminance information of imagined scenes were significantly less similar for comparisons of trials sampling similar directions compared to pairs sampling

dissimilar directions in a subset of participants. Specifically, Fisher z-transformed correlations were significantly lower (corrected for 24 comparisons) for image pairs from similar direction trials for twelve participants when considering the phase information (minimum significant T_{3838} = -3.25; maximum significant T_{3838} = -4.42) and for six participants when considering luminance information (minimum significant T_{3838} = -3.11; maximum significant T_{3838} = -4.70). Based on these results it seems highly unlikely that differences in visual similarity of the scenes to be imagined drive the observed increased representational similarity of trial pairs sampling similar directions in the parahippocampal gyrus.

Representational similarity analysis: six-fold symmetry

In rodents, grid cells are typically found in the medial entorhinal cortex (Hafting et al., 2005). Recently, two reports suggested the posterior medial part of the human entorhinal cortex as the human homologue of the rodent medial entorhinal cortex based on local and global connectivity patterns (Maass et al., 2015; Navarro Schröder et al., 2015). Therefore, we expected an influence of grid-like representations on imagination only in the posterior medial entorhinal cortex (pmEC). The anterior lateral part of the entorhinal cortex (alEC) served as a control region. ROI masks for the left and right pmEC and alEC from our previous report (Navarro Schröder et al., 2015) were first registered to the MNI template at a resolution of 2 mm using FSL (1 mm resolution for display in Figure 3e and Figure 3 - figure supplement 2a). For each participant, the masks were then warped to the anatomical scan with a resolution of 2 mm, which again served as the reference space for the different runs.

The analysis was performed on the same preprocessed and motion-corrected volumes as the one-fold symmetry analysis. Again, RSA was performed on the residuals of a GLM with the motion parameters as predictors. For each ROI, we calculated pattern similarity for all pairs of trials and defined two conditions to compare Fisher z-transformed Pearson correlations coefficients. We hypothesized that a possible influence of grid-cell representations on imagination should be reflected in a 60° periodicity of pmEC activity patterns (Figure 3ab and Figure 3 - figure supplement 1ab) analogously to the 60° modulation of the amplitude of the BOLD signal observed during navigation (Doeller et al., 2010). Extending this previous method, we pursued a more parsimonious and more powerful multivariate approach, which does not rely on the univariate estimation of a putative grid orientation on an independent dataset (which is only feasible with exhaustive sampling of directions during navigation (Doeller et al., 2010)). Here, pattern similarity in pmEC was

expected to be modulated by the angular difference of the directions sampled in a pair of trials (Figure 3c and Figure 3 - figure supplement 1c). To test this hypothesis two conditions were defined based on the modulus 60° of the angular difference between the sampled directions in a trial pair (Figure 3d). This remainder of modulus 60° could either be 0° or 30° due to the regular sampling of directions in our paradigm. We calculated the difference in mean pattern similarity between the 0° modulo 60° and the 30° modulo 60° condition for each subject. On the group level, we tested for increased pattern similarity in the 0° modulo 60° condition by comparing these pattern similarity difference values against 0 using a one-tailed one-sample T-test. Reported p-values are Bonferroni corrected for separate comparisons in left and right pmEC. To visualize whether the significant difference observed between the conditions in pmEC reflected a consistent 60° modulation of pattern similarity across angular differences, we plotted pattern similarity as a function of angular difference between trials pairs (Figure 3 - figure supplement 4).

Permutation-based significance testing

To further assess the increased pattern similarity for trials from the 0° modulo 60° and the 30° modulo 60° condition in left pmEC, we additionally performed a permutation-based non-parametric significance test. We computed the probability of the observed pattern similarity difference between the conditions under a permutation-based null distribution for each participant. Null distributions were obtained by shuffling the trial labels and analyzing the resulting correlation matrix for each of 10,000 permutations. The resulting p-value was converted to a z-statistic (allowing both positive and negative values). On the group level, we used a one-sample t-test (non-parametric with 10,000 permutations) to test these z-statistics against 0.

Grid-like entorhinal signals during imagination: Control analyses

As can be seen in a Figure 1a, the buildings relevant for the direction imagination task were distributed in a hexagonal pattern across Donderstown. A series of control analyses described in detail below suggests that the 60° modulation of pattern similarity in pmEC was not imposed by the layout of Donderstown and the specifics of the task design (Figure 3 - figure supplement 6 and Figure 3 - figure supplement 7).

Exclusion of trial pairs with the same start location. The number of comparisons of trials with the same start location was different between the 0° modulo 60° and the 30° modulo 60° condition. Thus, the 60° modulation of pattern similarity could have potentially been driven by the number of trial pairs in which the same start building had to be imagined. To account for this possibility, we excluded all pairs consisting

of trials with the same start location in a first control analyses (Figure 3 - figure supplement 6, bar I).
Exclusion of trial pairs with the same target location. With respect to the target locations, a similar aspect was considered. Due to the sampling of directions from the start locations near the center of Donderstown, buildings could only serve as targets for trials sampling a limited range of directions. Since trial pairs with the same target location were more likely to belong to the 0° modulo 60° condition, we excluded comparisons of trials with the same target location to rule out that the 60° modulation of pattern similarity in pmEC was driven by more frequent imagination of the same target building (Figure 3 - figure supplement 6, bar II).

Exclusion of trial pairs with the same combination of start and target location. A third control analysis excluded pairs of trials with the same start and target building combination to ensure that imagination of identical scenes did not drive the effect (Figure 3 - figure supplement 6, bar III).

Exclusion of trial pairs from the same MRI run. To rule out a potential influence of temporal autocorrelation, we restricted the analysis to pairs of trials from different task blocks. The effect also remained significant when excluding all comparisons of trials from the same block (Figure 3 - figure supplement 6, bar IV).

Exclusion of trial pairs with target locations in inner ring of buildings. In some trials, directions were sampled using target buildings located on the inner ring of buildings (see Figure 1 - figure supplement 2). Average angular error was lower for trials targeting a building on the inner compared to the outer ring (T_{23} = -5.29, $p < 0.001$). Pairs in which both trials sampled directions using target locations on the inner ring made up 12.23% of comparisons in the 0° modulo 60° condition and did not occur in the 30° modulo 60° condition. When excluding these comparisons from the analysis, pattern similarity remained increased for the 0° modulo 60° condition compared to the 30° modulo 60° condition (Figure 3 - figure supplement 6, bar V).

Rationale for controlling distance-related effects. In a second set of control analyses we additionally considered the Euclidean distance between the start and the target location as described above for the absolute directional coding effect. Again we adopted three different approaches to combine the distance information for a given trial pair (Figure 2 - figure supplement 2a) and computed a pairwise distance matrix for each measure. We assessed whether a distance measure differed between the 0° modulo 60° and the 30° modulo 60° condition by testing the difference between

the mean distance values of the two conditions against 0 using a one-sample t-test, analogous to our approach in the pattern similarity analysis. The results of these tests are reported in the following paragraphs. In each control analysis, we then computed a GLM with a continuous distance predictor and pairwise Fisher z-transformed correlation coefficients as the criterion. The residuals of these GLMs (i.e. pattern similarity that could not be explained by the distance predictor) were then used to compute the pattern similarity differences between the 0° modulo 60° and the 30° modulo 60° condition (Figure 3 - figure supplement 7). After all analyses, pattern similarity remained significantly greater for trial pairs in the 0° modulo 60° condition in the left pmEC. The effect also remained significant when using binary (high vs. low) distance predictors (all $T_{23} > 2.44$, all $p < 0.03$).

Controlling for mean distance of trial pairs. First, we averaged the lengths of the two vectors. The mean Euclidean distance was higher for pairs in the 0° modulo 60° condition ($T_{23} = 5.25$, $p < 0.001$). By accounting for the mean length of the two vectors, we exclude a potential influence of imagined distance to the target location on pattern similarity (Figure 3 - figure supplement 7, bar I).

Controlling for different distances within trial pairs. A second possibility could be a modulation of pattern similarity based on the difference in length of the trials defining a pair. Trial pairs from the 0° modulo 60° condition had more similar distances ($T_{23} = -57.02$, $p < 0.001$). Potentially, pattern similarity could be increased for trials of similar length. Therefore, we used the difference in length of the two trial vectors in a pair as a predictor (Figure 3 - figure supplement 7, bar II).

Controlling for the 'neighborhood distance'. Lastly, we quantified the distance of the six vectors connecting all four buildings in a trial pair. This resulted in a 'neighborhood' distance, which was low when both start and both target buildings were in the same area of Donderstown. These distances were larger for pairs in the 0° modulo 60° condition ($T_{23} = 54.20$, $p < 0.001$). We attempted to control for effects of imagining directions in specific parts of Donderstown with this analysis (Figure 3 - figure supplement 7, bar III).

Representations of cardinal directions. To examine whether potential representations of the cardinal directions were detectable in the pattern similarity structure exhibited by the pmEC during the direction imagination task, we compared pattern similarity values of pairs where both trials sampled a cardinal direction against pairs where this was not the case (Figure 3 - figure supplement 9a). The cardinal directions

were defined based on the orientation of Donderstown as displayed for example in Figure 1a. Pattern similarity was not increased for cardinal direction pairs in either pmEC (left: T_{23} = -0.136, p = 0.893; right: T_{23} = -0.449, p = 0.658, Figure 3 - figure supplement 9b) or alEC (left: T_{23} = 0.266, p = 0.793; right: T_{23} = 0.530, p = 0.601).

To ascertain that trial pairs sampling directions using buildings located on the same street running along the North-South axis of Donderstown (based on its orientation for example in Figure 1a), which could possibly bias participants' representations of the city due to a number of streets roughly aligned with it, were not driving our effects, we excluded all comparisons of pairs in which both trials sampled directions along this street from the analysis (see Figure 1 - supplement 2). Pattern similarity was significantly greater for pairs of trials sampling similar directions compared to pairs with larger angular differences (T_{23} = 4.00, p < 0.001) in the peak voxel of the parahippocampal cluster from the main absolute directional coding analysis (Figure 2), indicating that absolute directional coding was independent of pairs of trials sampling directions along the same street running along the North-South axis. Similarly, when examining the 60° modulation of pattern similarity values in left pmEC after excluding all trial pairs sampling along this street, the difference between the conditions remained significant (T_{23} = 2.40, p = 0.025).

A number of considerations further make an over-representation of the North-South axis due to directions sampled on a street running along this axis unlikely. Participants' performance did not differ between trials sampling along the North-South axis compared to other directions. We did not observe differences in accuracy (mean ± standard deviation (SD) for average absolute angular errors of trials sampling North-South 35.9° ± 24.5° vs. 33.2° ± 18.5° for other directions; means not significantly different: T_{23} = 1.06, p = 0.300) or consistency (mean ± SD of individual standard deviations for absolute angular errors for North-South trials 31.2° ± 19.5° vs. 30.1° ± 13.5° for other directions; standard deviations not significantly different: T_{23} = 0.54, p = 0.596) of participants' direction estimates. Furthermore, participants were recruited from the student body of a Dutch university and were therefore most familiar with irregular European street layouts (unlike the rectangular grid network common in the US), in which the use of cardinal directions is less common because most streets are curved and do not follow the cardinal directions. Importantly, even if participants preferentially represented directions along the North-South axis, which is unlikely based on the reasons laid out above, the population response of entorhinal grid cells would still remain the most likely explanation for the observed 60° modulation of pattern similarity values in pmEC,

which could not be explained by the presence of one cardinal axis.
Four-fold symmetry. To further corroborate the specificity of the 60° modulation observed in pmEC, we conducted a control analysis testing for a 90° modulation of pattern similarity values. A four-fold similarity pattern would be inconsistent with grid-cell like representations (Doeller et al., 2010; Kunz et al., 2015). We compared pattern similarity as function of angular difference for trial pairs with a 0° modulo 90° against trials pairs with 30° or 60° modulo 90° (Figure 3 - figure supplement 9c). There was no significant 90° modulation of pattern similarity in either pmEC (left: T_{23} = -0.48, p = 0.637; right: T_{23} = -1.81, p = 0.084, Figure 3 - figure supplement 9d) or alEC (left: T_{23} = -0.83, p = 0.413; right: T_{23} = 0.50, p = 0.618).

Estimation of signal-to-noise ratio in the entorhinal cortex
In order to compare signal quality in the entorhinal ROIs we calculated the temporal signal-to-noise ratio (tSNR). We quantified tSNR as the mean signal within a region divided by the standard deviation of this signal over time. A two-way repeated measures ANOVA with the factors region (pmEC vs. alEC) and hemisphere (left vs. right) revealed no significant differences in tSNR (no main effect of region ($F_{1,23}$ = 0.60, p = 0.448), no main effect of hemisphere ($F_{1,23}$ = 0.00, p = 0.953) and no interaction ($F_{1,23}$ = 0.97, p = 0.336); Figure 3 - figure supplement 3).

Whole-brain searchlight analyses
Our main analysis and subsequent control analyses described above focused on the head direction network and pmEC, respectively, based on our specific a priori hypotheses and previous studies (Baumann and Mattingley, 2010; Chadwick et al., 2015; Doeller et al., 2010; Marchette et al., 2014; Vass and Epstein, 2013, 2016). In particular, our pmEC ROI was motivated anatomically by the presence of grid cells in the rodent medial entorhinal cortex (Hafting et al., 2005) and by the functional identification of its human homologue region in our previous work (Navarro Schröder et al., 2015). Due to our complex novel imagination paradigm, set in a realistic large-scale VR city, we opted to further explore our data. We performed exploratory whole-brain searchlight analyses with the same parameters as reported above to test whether additional brain regions might display absolute directional coding and a 60° modulation in their pattern similarity structure, respectively. No brain regions other than the parahippocampal gyrus were observed for the absolute directional effect at a significance threshold of $p < 0.001$ uncorrected and a voxel extent threshold of ten voxels. Results for the six-fold symmetry whole-brain searchlight analysis are shown in Figure 3 - figure supplement 10. We did not perform searchlight analyses within pmEC or alEC due to the small size of these entorhinal subregions.

Exploration of behavioral data

We investigated whether a 60° modulation of absolute angular error values was also present in the behavioral data. We multiplied angular error values to combine performance for all pairs of trials and compared the resulting values as a function of angular difference of the directions sampled in a pair. Analogous to the 60° modulation analysis, we calculated the difference between combined error values in the 0° modulo 60° and the 30° modulo 60° condition for each participant on the first level and tested for a difference on the second level using a one-sample t-test. The combined error values were not different between the conditions (T_{23} = 1.24, p = 0.227, Figure 3 - figure supplement 8).

We further explored the behavioral data obtained during the direction imagination task. To assess whether the distance between the start and target building of a trial was predictive of the angular error of that trial, we calculated Pearson correlations between the absolute angular error and the distance from start to target building for each participant. This relationship did not reach statistical significance in any of the participants (mean r = -0.08 ± 0.09 standard deviation, range -0.24 - 0.10; all Bonferroni-adjusted p > 0.454).

FIGURE SUPPLEMENTS

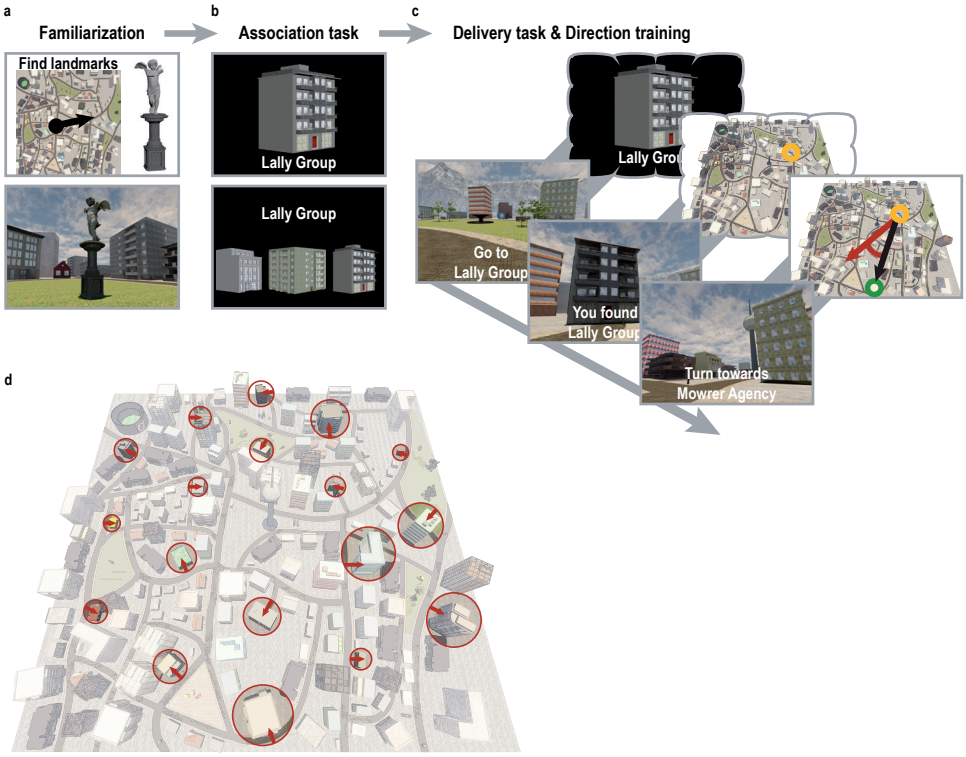


Figure 1 - figure supplement 1 | Overview of behavioral training.

(**a**) To familiarize themselves with the controls of the computer game and the layout of the city, participants explored Donderstown for 10 minutes and searched for a set of landmarks irrelevant for the direction imagination task. This exploration phase was omitted in the second training session. The black circle and arrow on the map in the top panel indicate the participants' position and orientation when first encountering Donderstown. (**b**) Subsequently, participants learned the names of 18 task-relevant buildings (top) to criterion. Knowledge of the building names was assessed in test blocks during which participants had to select the building belonging to the presented name from a display of three buildings by pressing one of three buttons (bottom). (**c**) For the remainder of the session, participants were trained on the building locations in Donderstown. Bottom row shows the trial structure as presented to the participants, top row for illustration only. Participants were instructed to navigate to the building whose name was presented on the screen. Once the building was located, participants encoded the position and were then asked to estimate the direction to the following target building. Performance was measured as the number of buildings located during the training session and the absolute angular difference between the estimated direction and the correct direction as defined by the current location and the new target building. (**d**) Overview of Donderstown highlighting the task-relevant buildings, which largely differed in features salient from the first-person perspective such as size, shape and rotation with respect to the hexagonal building layout (red arrows), which makes an influence of the regular arrangement of their entrances on participants' cognitive representation of the city unlikely.

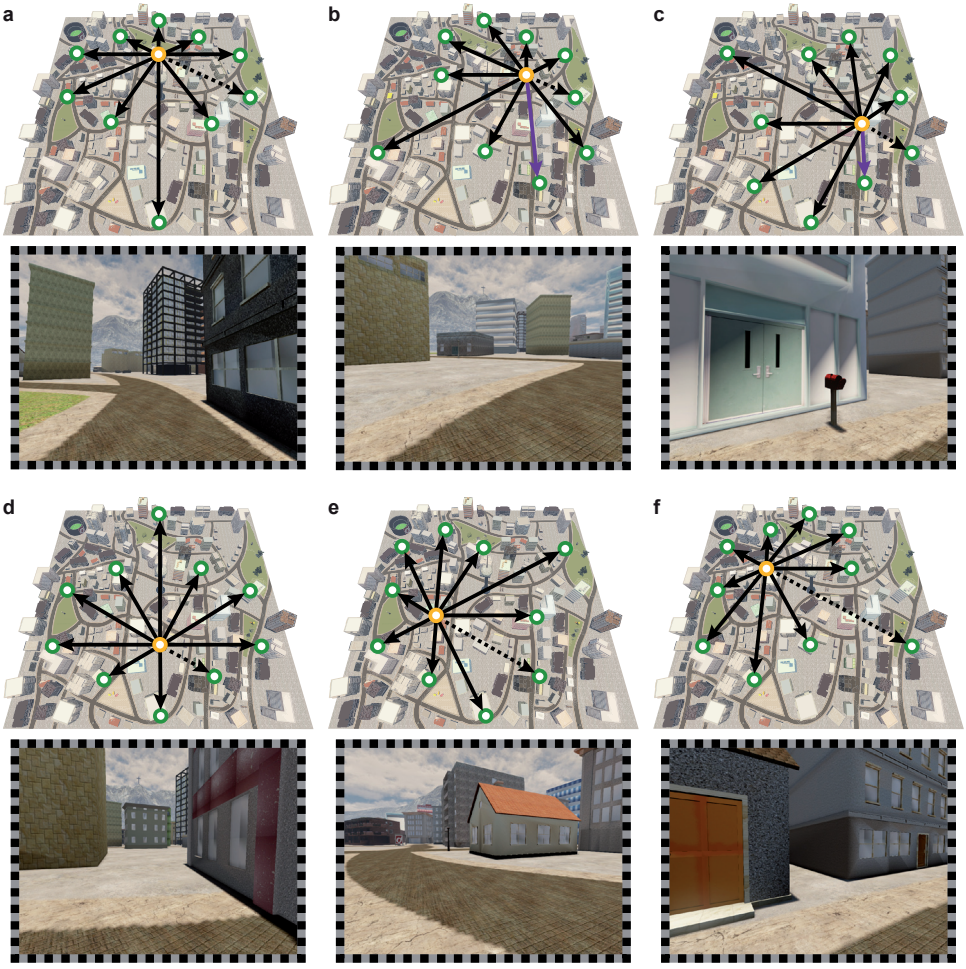


Figure 1 - figure supplement 2 | Sampling of directions in the imagination task.

(**a-f**) From each of the six start locations (yellow circles) ten directions were sampled. Directions (black arrows) were defined based on the angle of the vector connecting the start and the target locations (green circles). Screenshots show view from Donderstown corresponding to direction indicated by dashed arrow. Note that start locations could also be goal locations. The building combinations used in the direction imagination task were carefully counterbalanced so that in a task block of 24 trials each direction was sampled twice, each start building served as a start location four times and each building combination did not occur more than twice throughout the experiment (see Materials and methods). Trials sampling directions using buildings located on the same street (purple in **b** and **c**) were subject to an additional control analysis (see Materials and methods).

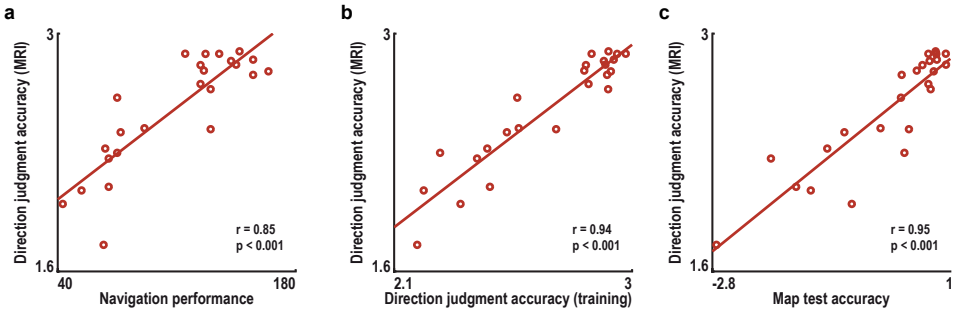


Figure 1 - figure supplement 3 | Accuracy of direction judgments during imagination task is related to behavioral performance during training and a post-scan map test.

(**a**) Across subjects, the accuracy of direction judgments in the direction imagination task (ϖ minus the mean angular difference between the correct and indicated directions in radians) correlated significantly with navigational success indexed by the number of buildings found during training. (**b**) Accuracy of direction judgments during the training sessions also correlated highly with performance during the imagination task. (**c**) Additionally, performance during the imagination task was correlated with z-scored accuracy in the post-scan map test. All correlations remained highly significant when controlling for variability in the time spent navigating the VR city during training using partial correlations (partial correlations coefficients $r > 0.84$, all $p < 0.001$). Correlations were also significant when calculated between the accuracy of direction judgments and training performance measures separately for the two training sessions (all correlations coefficients $r > 0.80$, all $p < 0.001$).

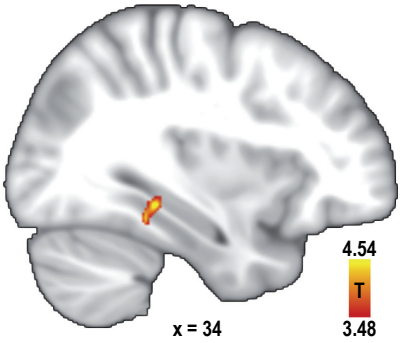


Figure 2 - figure supplement 1 | Increased pattern similarity for similar directions after excluding trial pairs sampling a direction with the same combination of buildings.

To control for the possibility of increased pattern similarity in the parahippocampal gyrus for similar direction pairs being due to the imagination of identical scenes, we excluded trial pairs from the analysis in which the same combination of start and target building was used. This revealed a significant cluster of voxels (peak voxel MNI coordinates: 34 -36 -10, $T_{23} = 4.54$, $p = 0.042$ corrected for multiple comparisons using small volume correction) very similar to the one observed in the main analysis (see Figure 2b). For display purposes, the statistical map is thresholded at $p < 0.001$ uncorrected.

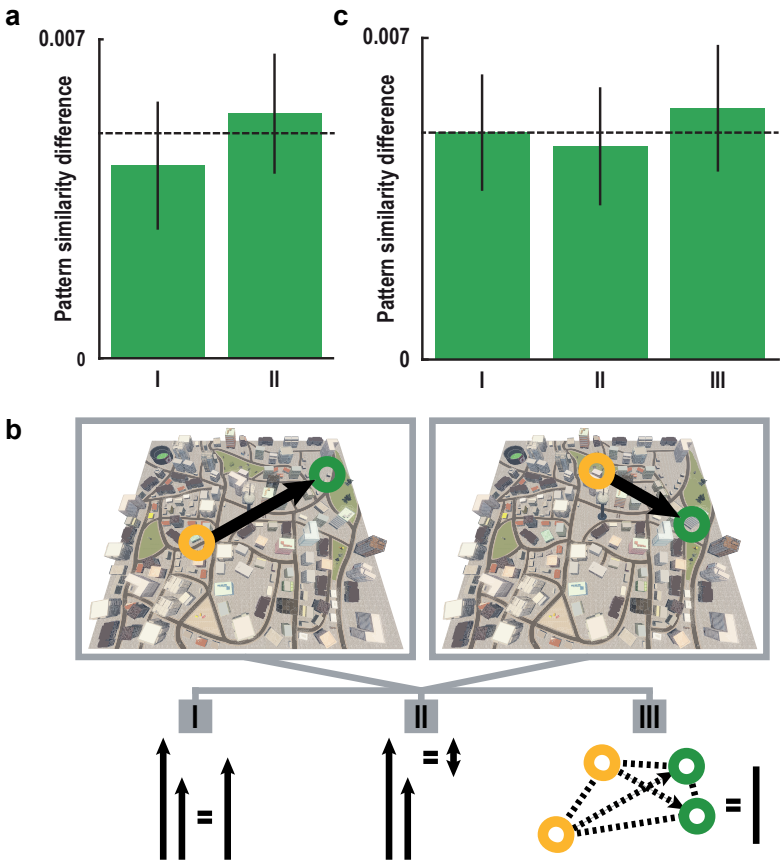


Figure 2 - figure supplement 2 | Absolute directional coding during imagination is independent of locations and distances in Donderstown.

(**a**) Pattern similarity difference between trial pairs sampling similar and dissimilar directions in the peak voxel of the main absolute directional coding analysis (Figure 2) after exclusion of comparisons with trials using the same start (bar I) and the same target location (bar II). Both $T_{23} > 3.00$, both $p < 0.007$. (**b**) We considered the distances between start and target locations in a trial pair by controlling for three distance measures, which differed between pairs of trials sampling similar and dissimilar directions. The distance measures are illustrated based on two example trials. For each trial pair, we calculated (I) the mean length of the vectors connecting start and target location, (II) the difference in length of the two vectors and (III) the mean length of the vectors connecting all four relevant locations of a pair. In separate GLMs we used the distance measures as predictors of pairwise pattern similarity. (**c**) Mean pattern similarity difference in peak voxel of cluster from main absolute directional coding analysis (Figure 2) between trial pairs sampling similar directions and pairs sampling dissimilar directions computed on the residuals of the GLMs. With this approach we controlled for pattern similarity due to (I) the average distance from start to target location in a trial pair, (II) the difference in distance from start to target location in a trial pair and (III) the average distance between all four buildings in a trial pair (see Materials and methods). All $T_{23} > 3.60$, all $p < 0.001$. Bars in a and c show mean pattern similarity difference with error bars reflecting S.E.M., dashed line shows mean pattern similarity difference in peak voxel from main analysis.

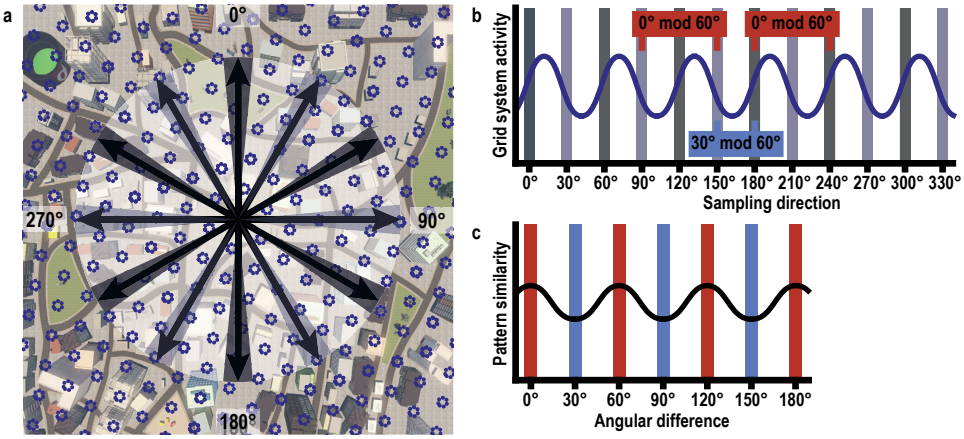


Figure 3 - figure supplement 1 | Rationale of 60° modulation analysis.
(**a**) Six-fold symmetric firing fields of a hypothetical grid cell (dark blue dotted circles) superimposed on a top-down view of Donderstown. Black arrows indicate the twelve sampled directions. Light and dark shading indicates directions (multiples of) 60° apart. (**b**) The firing rate of the hypothetical grid cell as a function of sampling direction exhibits a 60° modulation. Shading shows sampled directions with red and blue markers illustrating the two conditions. Note that the oscillatory firing pattern is sampled at the same phase in the 0° modulo 60° condition, but at different phases in the 30° modulo 60° condition. (**c**) Based on this, we expected increased pattern similarity when comparing trial pairs from the 0° modulo 60° condition to trial pairs from the 30° modulo 60° condition. The difference between the conditions is smaller than in Figure 3c due to the different sampling of directions with respect to the grid orientation.

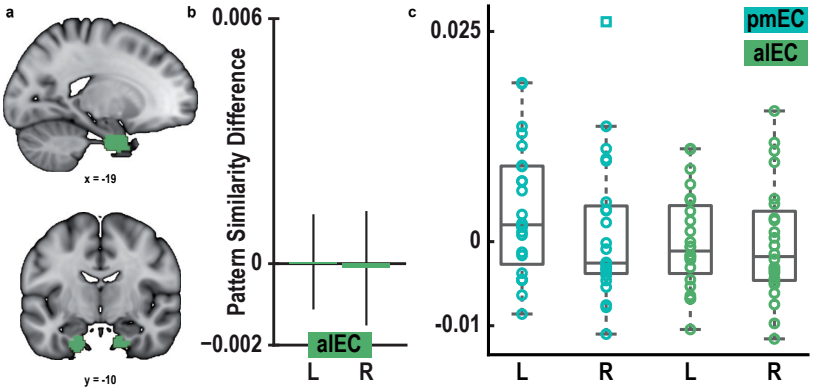


Figure 3 - figure supplement 2 | Pattern similarity difference between 0° modulo 60° condition and 30° modulo 60° condition in pmEC and alEC.

(**a**) ROI mask for anterior lateral entorhinal cortex (alEC) based on our previous report (Navarro Schröder et al., 2015). (**b**) Pattern similarity did not differ between the 0° modulo 60° and the 30° modulo 60° condition in alEC (T_{23} = 0.04, p = 0.97 and T_{23} = -0.08, p = 0.94 for left and right alEC, respectively). (**c**) Colored markers show pattern similarity difference for each participant in pmEC and alEC. Boxplots indicate 25th and 75th percentile with the middle line representing median pattern similarity difference across participants. Whiskers extend to most extreme data points not considered outliers. Data points defined as outliers (values more than 1.5 times the interquartile range above the 75th percentile or more than 1.5 times the interquartile range below the 25th percentile) are represented by square markers.

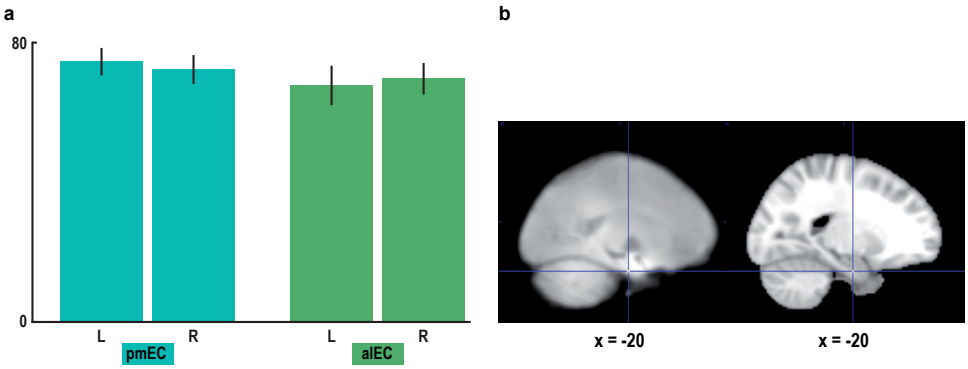


Figure 3 - figure supplement 3 I Signal quality in the entorhinal cortex.

(**a**) To assess signal quality in pmEC and alEC, we computed the temporal signal-to-noise ratio (tSNR, see Materials and methods). A repeated-measures ANOVA revealed neither a main effect of region ($F_{1,23}$ = 0.60, p = 0.448) or hemisphere ($F_{1,23}$ = 0.00, p = 0.953) nor an interaction between the factors region and hemisphere ($F_{1,23}$ = 0.97, p = 0.336). Bars represent mean and S.E.M. (**b**) Left slice shows the mean functional scan averaged across participants. The mean functional images from the four fMRI runs were averaged for each participant before averaging the resulting mean images across participants. Right slice shows the corresponding section of the MNI template. Note that for some participants the edge of the superior parietal lobe was outside the field of view.

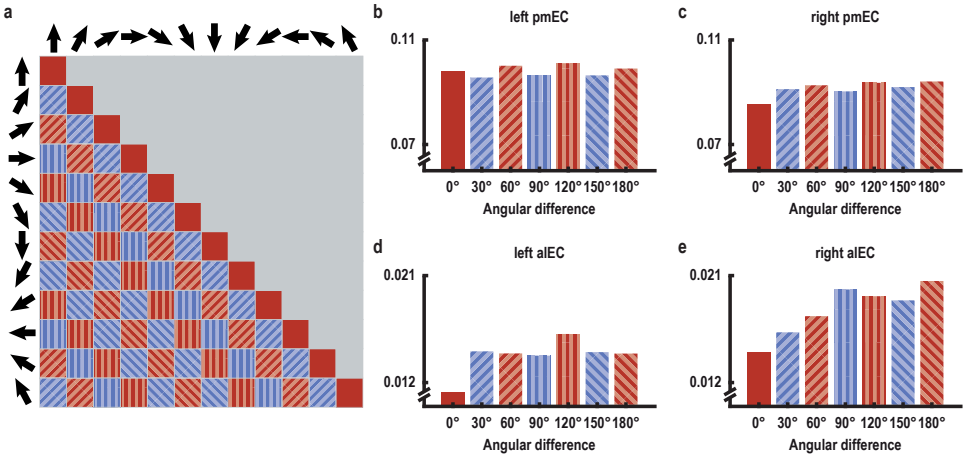


Figure 3 - figure supplement 4 I 60°-periodicity of pattern similarity is consistent across angular differences only in left posterior medial entorhinal cortex.

(**a**) Pattern similarity was analyzed based on the angular differences of the directions sampled in a trial pair. High pattern similarity was predicted for pairs in the 0° modulo 60° condition (red) in contrast to pairs in the 30° modulo 60° condition (blue). Filled pattern and color indicates angular differences and corresponds to bars in (**b-e**), which visualize average pattern similarity for all possible angular differences for exploratory purposes in the entorhinal ROIs. Note the consistent 60°-periodicity of the pattern similarity profile across angular differences in left pmEC. A statistical test was performed on the within-subject difference between the two conditions and was significant in left pmEC only (T_{23} = 2.37, p = 0.027, for all other ROIs, p > 0.5). Error bars for each angular difference would not reflect the statistical test performed and are therefore omitted.

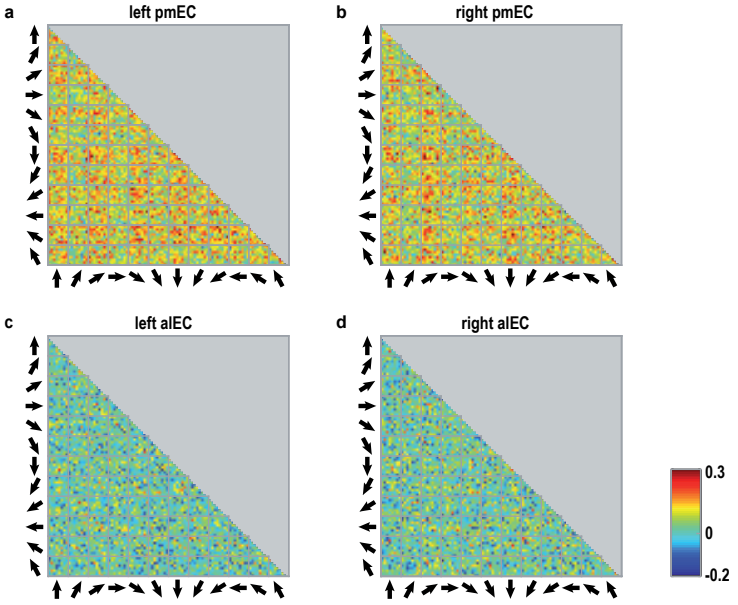


Figure 3 - figure supplement 5 | Pattern similarity structure across pair-wise comparisons of trials for entorhinal ROIs.

(**a-d**) Matrices show the pair-wise correlations between voxel patterns in the subregions of the entorhinal cortex (**a**: left pmEC; **b**: right pmEC; **c**: left alEC; **d**: right alEC) across all possible trial comparisons averaged across participants. Arrows signal sampled directions in a given pair of trials. Colorbar indicates size of the correlations in panels (**a-d**).

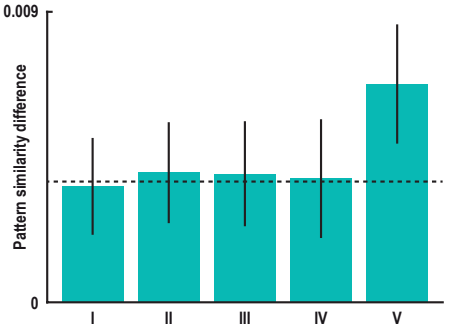


Figure 3 - figure supplement 6 | 60° modulation of pattern similarity during imagination is not driven by specifics of task design.

Pattern similarity difference between the 0° modulo 60° and the 30° modulo 60° condition remained significant in left pmEC after controlling for specifics of the design. Bars show the mean pattern similarity difference after excluding trial pairs with (I) the same start location (T_{23} = 2.39, p = 0.025), (II) the same target location (T_{23} = 2.57, p = 0.017), (III) the same combination of start and target location (T_{23} = 2.45, p = 0.022),(IV) pairs from the same run (T_{23} = 2.08, p = 0.049) and (V) pairs with target locations in the inner ring of buildings (T_{23} = 5.29, p < 0.001; see Materials and methods). This excludes potential influences of imagining the same start or same target location, the same combination of start and target location and temporal auto-correlations on the effect. Error bars indicate S.E.M., dashed line shows mean pattern similarity difference in left pmEC from main analysis (**Figure 3b**).

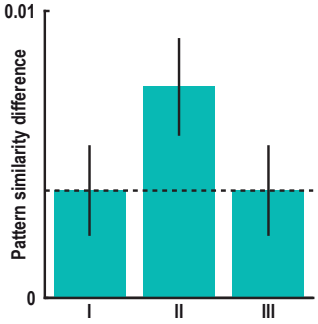


Figure 3 - figure supplement 7 I 60° modulation of pattern similarity during imagination after controlling for distance measures.

We controlled for the distances between start and target locations in a trial pair using three distance measures, which differed between the 0° modulo 60° and the 30° modulo 60° condition (see Materials and methods). The distance measures are illustrated in Figure 2 - figure supplement 2b. In separate GLMs we used the distance measures as predictors of pairwise pattern similarity and computed the mean pattern similarity difference between the 0° modulo 60° and the 30° modulo 60° condition on the residuals of these GLMs. With this approach we controlled for pattern similarity due to (I) the average distance from start to target location in a trial pair, (II) the difference in distance from start to target location in a trial pair and (III) the average distance between all four buildings in a trial pair (see Materials and methods). All $T_{23} > 2.36$, all $p < 0.03$, error bars indicate S.E.M., dashed line shows mean pattern similarity difference in left pmEC from main analysis (Figure 3b). The effect also remained significant when using binary (high vs. low) distance predictors (all $T_{23} > 2.44$, all $p < 0.03$).

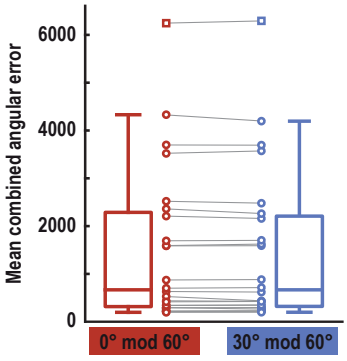


Figure 3 - figure supplement 8 I Behavioral performance for trial pairs in the 0° modulo 60° and the 30° modulo 60° condition.

Error values were multiplied for the two trials of each pair and averaged for the two conditions. Boxplots indicate 25th and 75th percentile with the middle line representing median combined error across participants. Whiskers extend to most extreme data points not considered outliers. Data points connected by lines show combined errors for all subjects in the two conditions (no difference between conditions, $T_{23} = 1.24$, $p = 0.227$). Data points defined as outliers (values more than 1.5 times the interquartile range above the 75th percentile or more than 1.5 times the interquartile range below the 25th percentile) are represented by square markers. Note that the participant shown here as an outlier performed above chance and that there were no outliers in our main pattern similarity analysis (see Figure 3 - figure supplement 2).

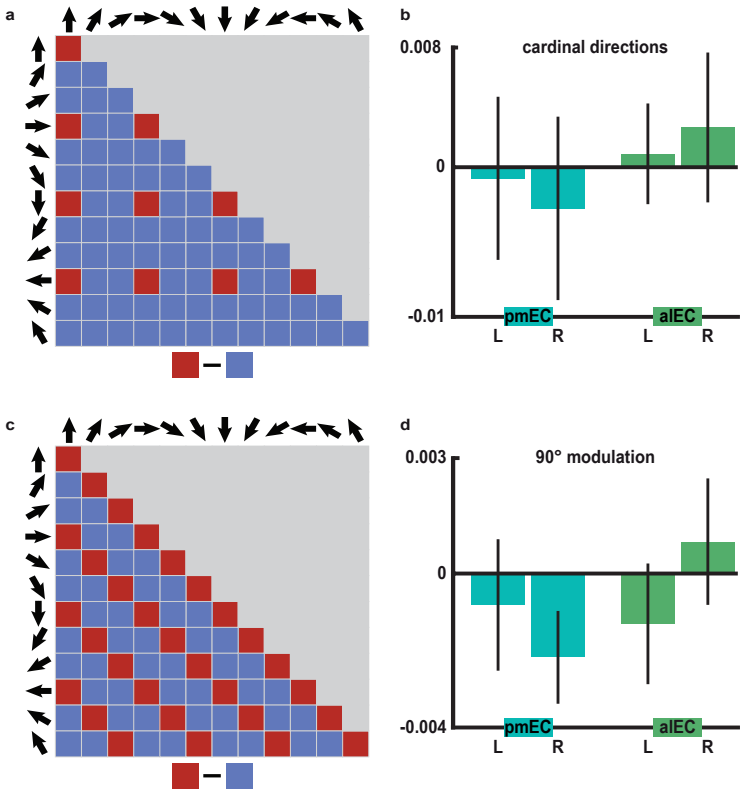


Figure 3 - figure supplement 9 | No evidence for representations of cardinal directions or 90° modulation of pattern similarity in the entorhinal cortex.

(**a**) Structure of pattern similarity values used to test for coding of cardinal directions in the entorhinal cortex. If entorhinal cortex activity would be sensitive to cardinal directions, high pattern similarity would be expected for pairs of trials sampling cardinal directions in comparison with trials sampling other directions. Note that for illustration purposes the tested similarity matrix is shown for comparisons across conditions, not single trials. (**b**) Pattern similarity did not differ between pairs of trials sampling cardinal directions and pairs of trials sampling other directions in pmEC (differences scores: left: T_{23} = -0.136, p = 0.893; right: T_{23} = -0.449, p = 0.658) or alEC (differences scores: left: T_{23} = 0.266, p = 0.793; right: T_{23} = 0.530, p = 0.601). (**c**) To corroborate the specificity of the 60° modulation of pattern similarity in pmEC, we examined a possible, yet biologically implausible four-fold symmetry in entorhinal pattern similarity values. We tested for increased pattern similarity for pairs of trials sampling directions 90° or multiples thereof apart (0° modulo 90° against 30° or 60° modulo 90°), using the same analysis logic as for the main analysis (0° modulo 60° against 30° modulo 60°) but now with a 90° periodicity. (**d**) Pattern similarity values did not differ between these conditions in pmEC (differences scores: left: T_{23} = -0.48, p = 0.637; right: T_{23} = -1.81, p = 0.084) or alEC (differences scores: left: T_{23} = -0.83, p = 0.413; right: T_{23} = 0.50, p = 0.618). Bars in (**b** and **d**) represent mean and S.E.M.

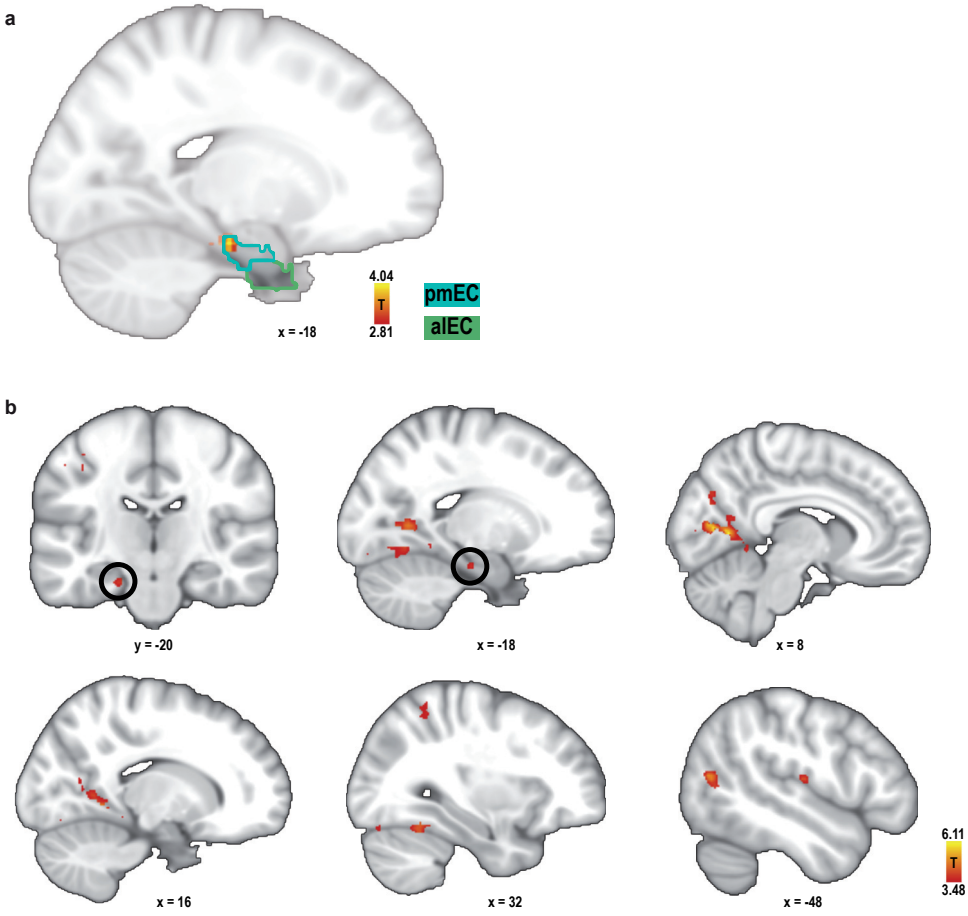
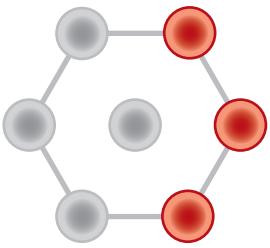


Figure 3 - figure supplement 10 I Searchlight analysis for 60° modulation of pattern similarity during imagination.

For each search sphere, the difference in pattern similarity for trial pairs in the 0° modulo 60° condition and the 30° modulo 60° condition was calculated. (**a**) One entorhinal cluster exhibited increased pattern similarity for the 0° modulo 60° condition (peak MNI coordinates: -18 -20 -22; T_{23} = 4.04, p = 0.046, FWE-corrected for multiple comparisons in bilateral pmEC and alEC using small volume correction) and is shown together with the masks of pmEC and alEC, outlined in dark and light green, respectively. This result confirms our finding from the ROI analysis. Statistical map is thresholded at p < 0.005 uncorrected and masked to show entorhinal cortex only. (**b**) Further exploratory whole-brain searchlight analysis showed greater pattern similarity for the 0° modulo 60° condition in the lingual gyrus (MNI coordinates: 8 -62 6, T_{23} = 6.11), precuneus (-20 -60 4, T_{23} = 5.35), cuneal cortex (10 -74 26, T_{23} = 5.26), lateral occipital cortex (-42 -66 6, T_{23} = 5.06), occipital pole (-32 -90 10, T_{23} = 4.93), supplementary motor cortex (0 -14 56, T_{23} = 4.54), central opercular cortex (-50 -16 12, T_{23} = 4.46), occipital fusiform gyrus (38 -74 -18, T_{23} = 4.36), angular gyrus (54 -56 26, T_{23} = 4.34), superior parietal lobule (30 -52 58, T_{23} = 4.12) and supramarginal gyrus (-64 -50 22, T_{23} = 4.09). Slices show the statistical image at 1 mm resolution at a significance threshold of p < 0.001. Labels were obtained from the Harvard-Oxford Structural Cortical Structural Atlas available in FSL. For each region peak voxel MNI coordinates and statistics are reported.

CHAPTER 3

Deforming the Metric of Cognitive Maps Distorts Memory



This chapter is accepted for publication:

J. L. S. Bellmund, W. de Cothi, T. A. Ruiter, M. Nau, C. Barry, C. F. Doeller (accepted for publication). Deforming the metric of cognitive maps distorts memory. *Nature Human Behaviour*. Preprint available at https://doi.org/10.1101/391201.

ABSTRACT

Entorhinal grid cells, characterized by spatially periodic activity patterns, are thought to provide a universal spatial metric. However, grid cell firing-patterns are distorted in highly polarized environments such as trapezoids. Additionally, the functional role of grid cells in guiding behavior remains elusive. Here, we leverage immersive virtual reality using a novel motion platform to test the impact of environmental geometry on spatial memory in participants navigating a trapezoid arena. Object position memory in the trapezoid was degraded compared to a square control environment. Consistent with grid pattern distortions in rodents, this effect was more pronounced in the narrow than the broad part of the trapezoid. Remarkably, even outside of the encoding environment, these distortions persistently affected both navigated and judged distance estimates of never experienced paths between remembered positions and reconstructed memory maps. These distorted memory maps in turn explained behavior better than objective maps. Our findings demonstrate that environmental geometry interacts with human spatial memory similarly to how it affects rodent grid cells – thus strengthening the putative link between grid cells and behavior as well as cognitive functions beyond navigation.

INTRODUCTION

The neural basis of navigation and spatial memory is well-studied and has resulted in the identification of functionally defined cell types encoding spatial information relevant to wayfinding (Moser et al., 2017; Epstein et al., 2017). Among these are grid cells, first identified in the entorhinal cortex of freely moving rodents, which exhibit six-fold periodic (hexadirectional) spatial firing extending across the environment (Hafting et al., 2005). This pattern can be described in terms of its scale, as well as its offset and orientation relative to the environment (Hafting et al., 2005; Moser et al., 2017). Along the dorso-ventral axis of medial entorhinal cortex , grid cells sharing similar spacing and orientations are organized in discrete modules (Barry et al., 2007; Brun et al., 2008; Stensola et al., 2012). Grid cells have been directly recorded in human patients undergoing pre-surgical screening (Jacobs et al., 2013; Nadasdy et al., 2017) and in human fMRI studies hexadirectional signals serve as a proxy measure for activity of the entorhinal grid system (Doeller et al., 2010; Kunz et al., 2015; Bellmund et al., 2016; Constantinescu et al., 2016; Horner et al., 2016; Navarro Schröder et al., 2017; Nau et al., 2018a; Julian et al., 2018; Stangl et al., 2018).

Theoretical work suggests that regular grid patterns provide a compact code for self-localization and function as a metric for space, supporting path integration and vector-based navigation (Moser et al., 2017; Hafting et al., 2005; McNaughton et al., 2006; Fiete et al., 2008; Burak and Fiete, 2009; Mathis et al., 2012; Bush et al., 2015; Herz et al., 2017; Banino et al., 2018). Thus, location is encoded by the conjunction of spatial phases across different modules – the population phase (Bush et al., 2015; Carpenter and Barry, 2016) – while the distance and direction between points can be derived from the relative difference in population phase (Bush et al., 2015). Yet direct empirical evidence demonstrating the behavioral relevance of grid cells remains scarce. However, in line with the proposed role of grid cells in spatial navigation and memory, the strength and coherence of hexadirectional signals across voxels in human entorhinal cortex predict the accuracy of subjects' spatial memory responses in virtual reality (VR) navigation tasks (Doeller et al., 2010; Kunz et al., 2015).

Despite their regular firing patterns (Hafting et al., 2005) and the consistent behavior of grid cell ensembles across different environments (Fyhn et al., 2007; Yoon et al., 2013) and behavioral states (Gardner et al., 2017; Trettel et al., 2017), environmental geometry strongly influences grid firing patterns (Krupic et al., 2015;

Stensola et al., 2015; Sun et al., 2015; Krupic et al., 2018). Thus, changes made to the geometry of a familiar enclosure produce commensurate changes to the scale of grid-patterns, resulting in differential rates of change in population phase for travel in the changed and unchanged dimension (Barry et al., 2007; Stensola et al., 2012). Similar manipulations made while humans navigate in VR environments produce complementary deficits in path integration (Chen et al., 2015). Equally, the configuration of static environments also influences grid regularity, for example, in square enclosures grid firing tends to orient relative to enclosure walls (Krupic et al., 2015; Stensola et al., 2015; Sun et al., 2015). More strikingly, in highly polarized enclosures such as trapezoids, grid-patterns are highly distorted and less regular than in control enclosures (Krupic et al., 2015). These changes were found to be especially pronounced in the narrow part of the trapezoidal enclosure with reduced symmetry, larger firing fields, increased spacing and a change of grid orientation – changes which do not appear to attenuate with continued exposure (Krupic et al., 2015). Similarly, in a quadrilateral environment with one slanted wall, firing fields of grid cells were observed to be consistently shifted away from the slanted wall, resulting in a local distortion of the grid (Krupic et al., 2018). Decoding of positions from simulated grid cells capturing this distortion was less accurate near the slanted wall (Krupic et al., 2018). However, the behavioral consequences of degraded positional information due to a compromised grid pattern, which might result in uncertainties about locations and distances in space (Carpenter and Barry, 2016; Krupic et al., 2015, 2018) (Figure 1), have not been explored. Here, we use environmental manipulations to shed light on how grid cell-based computations might support behavior.

Inspired by the distortions of grid-cell firing patterns in a trapezoidal enclosure in rodents, we investigated the effects of environmental geometry on human spatial memory. We hypothesized that environmental geometry-dependent distortions of the spatial metric provided by grid cell firing would affect mnemonic performance. Specifically, we predicted positional memory to be impaired in a trapezoidal compared to a square control environment. Relatedly, we expected that within the trapezoid performance should be more degraded in the narrow than the broad part of the enclosure. Importantly, object locations encoded using a distorted metric should further persistently affect subsequent spatial computations performed on the distorted memories. We therefore tested whether mnemonic distortions would persist outside of the trapezoidal environment and predicted differential estimates of identical distances between positions located in the narrow and broad part of the trapezoid, respectively.

RESULTS

Positional memory

We employed immersive VR to investigate effects of environmental geometry on human spatial memory (Figure 1). Wearing a head-mounted display, participants navigated a trapezoid and a square environment of equal surface area using a motion platform translating real-world rotations and steps into virtual movement. In both environments, participants learned the positions of six objects, organized in two triplets with matched inter-object distances in both halves of an environment (Figure 1B). Participants were tested on the positions of the objects after an initial learning phase by having to navigate to the remembered position of a cued object in each trial (Figure 1A). In a first step, we then compared positional memory, i.e. the Euclidean distance between this response position and the correct position, between the two environments. In line with degradations in grid cell firing patterns observed in rodents (Krupic et al., 2015), results showed larger response errors in the trapezoid than the square (Figure 2A, $t(36)=2.71$, $p=9.999\times10^{-5}$; bootstrap-based t-tests are reported throughout the manuscript, see Methods).

To ensure that this effect was not due to the fact that larger error distances are possible within the trapezoid because of its elongated shape, we calculated memory scores to account for differences in the distribution of possible errors for each position (Jacobs et al., 2016). In short, we generated a set of 1000 random locations uniformly covering the entire environment, compared the error distance of a given trial to the distances from the random locations to the trial's correct position and quantified the memory score as the proportion of random locations further away from the correct position than the observed response position. This yielded a score ranging from 0 (low memory) to 1 (perfect memory) for each trial, taking into account the range of possible errors based on the correct position and environmental geometry (Figure 2B). Importantly, memory scores were significantly lower in the trapezoid compared to the square (Figure 2C, $t(36)=-2.30$, $p=5.000\times10^{-4}$), ensuring that decreased positional memory was not due to different distributions of possible errors as a result of the elongated shape of the trapezoid.

In rodents, entorhinal grid-patterns are most distorted in the narrow part of the trapezoid enclosure (Krupic et al., 2015). Building on this observation, we predicted that positional memory should follow this pattern if such deformations are indeed detrimental to spatial memory. To account for the overall high performance levels of participants (Supplemental Figure 1A), we referenced memory performance in

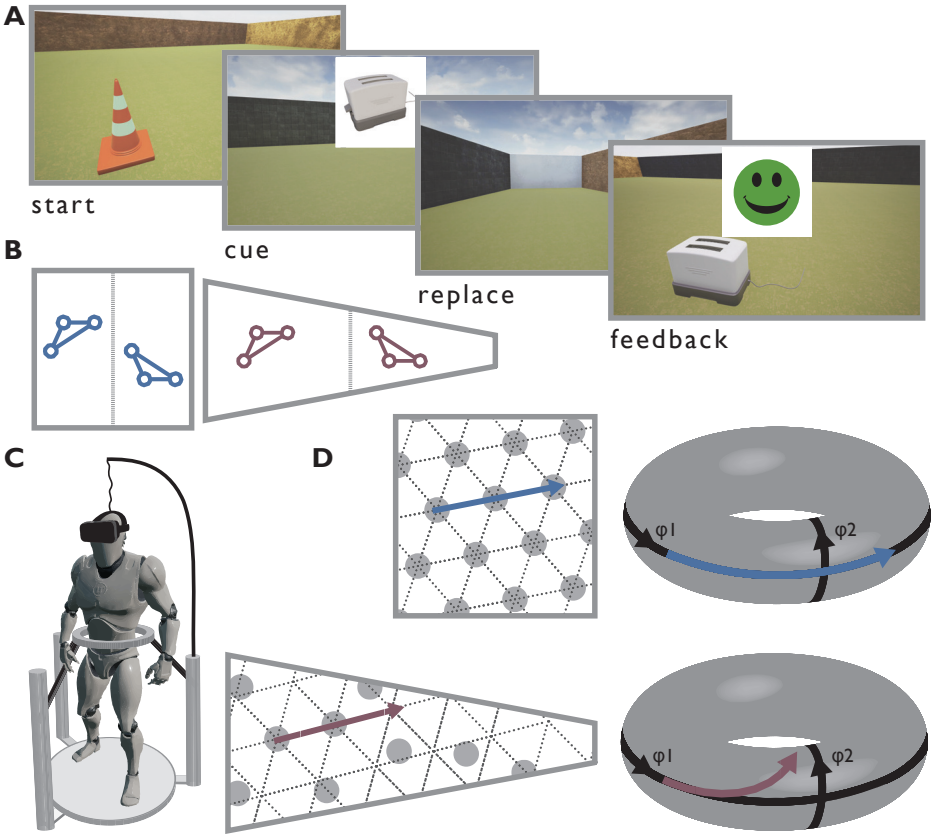


Figure 1 | Task design and rationale.

A. To commence a trial, participants walked to a start position marked by a pylon where they were cued with the image of an object. Subsequently, they navigated to the object's remembered position, which they indicated via button press, and received feedback. **B.** Circles illustrate an exemplary configuration of object positions. Two triplets of objects were positioned in each environment with one triplet in each half of each environment, yielding four triplets with matched distances between positions. **C.** Schematic of immersive VR setup with head mounted display and motion platform translating physical steps and rotations into virtual movement. **D.** Illustration of regular and distorted grid in square and trapezoidal environment, respectively. Dashed lines indicate axes of idealized grid and grey circles mark corresponding firing fields. The torus provides a continuous representation of the 2D-periodic population phase. Translational movement of a given distance in the square (blue arrow) results in a corresponding change of grid population phase (blue arrow on the torus). In the trapezoidal environment, a movement of the same Euclidean distance (purple arrow) results in a different change in population phase (purple arrow on the torus) due to the degraded grid pattern characterized by reduced symmetry as a result of changes in grid period and orientation. A similar effect is expected between the narrow compared to the broad part of the trapezoid (not shown) due to the more severe distortion of the grid pattern (see Methods).

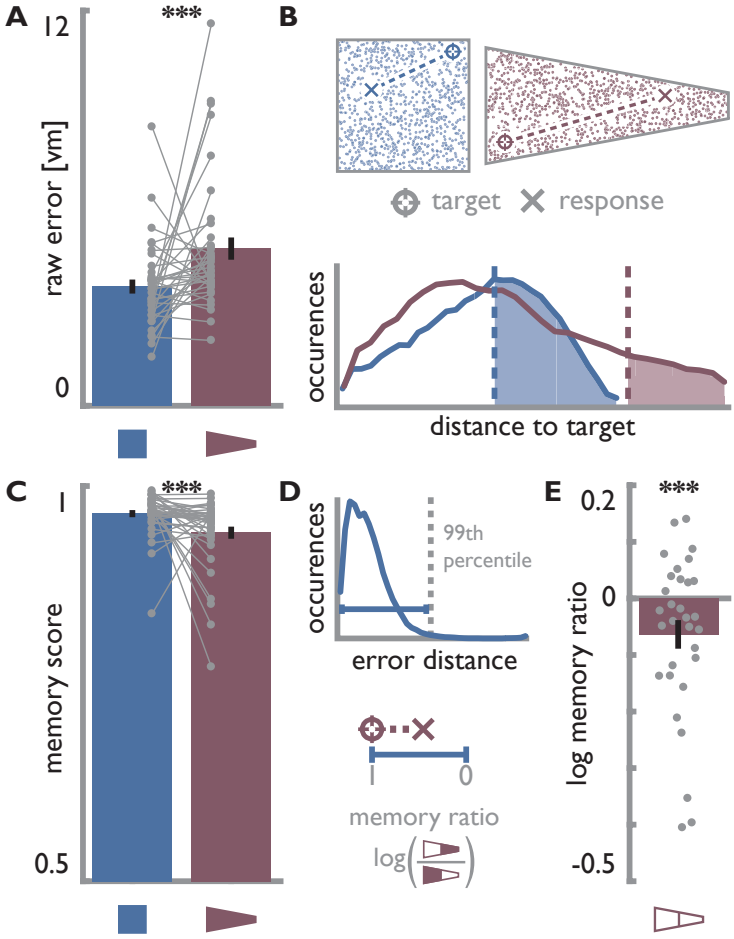


Figure 2 | Distorted positional memory in the trapezoid.

A. Raw replacement errors in virtual meters (vm) as measured by the distance between the correct and remembered position were larger for object positions in the trapezoid than in the square. **B.** Schematics illustrate the expression of positional memory using memory scores for two hypothetical trials. For both environments a distribution of random positions uniformly covering the available space was generated (top). For each trial, mnemonic performance was expressed as a memory score quantifying the proportion of random positions further away from the correct object position than the response location (bottom). **C.** Memory scores were lower in the trapezoid than in the square environment. Y-axis thresholded at chance level of 0.5. **D.** Schematics illustrate calculation of the memory ratio. Using the 99th percentile of error values across all participants and trials in the square environment as a reference (top), we quantified performance in the narrow and broad part of the trapezoid on a scale from 1 (perfect memory) to 0 (replacement error the same size or larger than the reference value obtained from performance in the square). The log memory ratio contrasts average performance in the narrow and broad part of the trapezoid. **E.** The memory ratio was significantly below 0, indicating lower positional memory in the narrow compared to the broad part of the trapezoid. 6 participants were excluded due to log memory ratios more than 1.5 times the interquartile range above or below the upper and lower quartile respectively. See Supplemental Figure 1B for full dataset. All bars show mean±SEM, grey circles indicate individual subject data with lines connecting data points from the same participant. *** $p<0.001$

the trapezoid to memory performance in the square to allow for a comparison of replacement errors between the two parts of the environment despite differences in local geometry. Specifically, we chose the 99th percentile of error distances across all trials and participants as a reference. Each trial was then assigned a score calculated as one minus the proportion of its replacement error of this reference value obtained from the square (Figure 2D). This resulted in memory scores, comparable across subjects, with larger values indicating better performance. To compare positional memory between the narrow and broad part of the trapezoid we calculated the log memory ratio to quantify the relation of mean memory scores in the two parts. This ratio was significantly below 0 (Figure 2E, $t(30)=-2.52$, $p=9.999\times10^{-5}$), indicating decreased memory performance in the narrow compared to the broad part of the environment. This was also true for the memory ratio calculated from the raw error scores (Supplemental Figure 1B). The log memory ratio of the trapezoid was lower than the 5th percentile of a surrogate distribution of memory ratios obtained from comparing positional memory between the halves of the square, indicating a significant difference between the environments (Supplemental Figure 1C, $Z=-2.48$, $p=0.007$). Taken together, the profile of positional memory observed is in line with our predictions derived from deformations of grid cell firing patterns with degraded positional memory in the trapezoidal compared to the square environment and more impaired performance in the narrow than the broad part of the trapezoid.

These differences in positional memory were not due to differential navigation behavior: The excess path length of participants' navigation paths from the start position of a given trial to where they remembered the object did not differ in the trapezoid compared to the square environment or between the two parts of the trapezoid (Supplemental Figure 2, square vs. trapezoid: $t(36)=-0.95$, $p=0.144$; broad vs. narrow trapezoid: $t(36)=-0.11$, $p=0.865$). Further, walking speeds did not differ between the two environments or the sub-parts of the trapezoid (Supplemental Figure 2, trapezoid vs. square: $t(36)=-0.01$, $p=0.973$; broad vs. narrow trapezoid: $t(36)=1.15$, $p=0.079$). Euclidean distances from the start to the correct object positions were not related to spatial memory performance (Supplemental Figure 2, square: $t(36)=0.17$, $p=0.863$, trapezoid: $t(36)=0.17$, $p=0.869$, trapezoid broad: $t(36)=1.37$, $p=0.177$, trapezoid narrow: $t(36)=-0.01$, $p=0.998$).

To explore participants' navigation behavior in more detail we next examined their body and head orientation during the replacement period relative to the direction from start to response position in each trial (Supplemental Figure 3). Both body and head orientation of participants were significantly clustered around this direction

in square (body: v=36.84, p=1.086×10^{-17}; head: v=36.68, p=1.500×10^{-17}) and trapezoid (body: v=36.89, p=9.738×10^{-18}; head: v=36.68, p=1.490×10^{-17}) and the distributions of mean orientations were not different between the two environments (body: F(1,72)=0.02, p=0.889; head: F(1,72)=0.14, p=0.709). Similar results were obtained for the body and head orientations when comparing trials targeting objects in the broad and narrow parts of the trapezoid (body broad: v=36.73, p=1.357×10^{-17}; body narrow: v=36.73, p=1.350×10^{-17}; difference body: F(1,72)=1.53, p=0.220; head broad: v=36.54, p=1.957×10^{-17}; head narrow: v=36.65, p=1.579×10^{-17}; difference head F(1,72)=0.05, p=0.830). Hence, we do not expect average facing direction to influence our key comparisons between the two environments or within the trapezoid. The circular variance around each trial's average body direction did not differ between environments (trapezoid vs. square: t(36)=1.06, p=0.118) or the trapezoid parts (narrow vs. broad: t(36)=1.14, p=0.076), but the circular variance of participants' facing directions was greater in the trapezoid than in the square (t(36)=2.57, p=9.999×10^{-5}) and greater in the narrow than broad part of the trapezoid (t(36)=2.13, p=5.000×10^{-4}), potentially reflecting increased visual exploration and orientation to compensate for impaired spatial memory. Do attentional resources and task demands differ between test environments? This seems unlikely as our design included a secondary task in which participants memorized color changes of an extramaze cue and later estimated durations between color change events (see Methods). Mean (t(36)=-0.10, p=0.873) and absolute (t(36)=-0.32, p=0.629) estimation errors as well as error variability (t(36)=-0.81, p=0.205) did not differ between square and trapezoid (Supplemental Figure 6), suggesting comparable attentional resources were available in both environments.

Mnemonic distortions outside of the environment

Next, we sought to address whether the impact of environmental geometry on spatial memory persisted outside of the learning environment. To this end, we asked participants to estimate distances between the positions of object pairs in two modalities subsequent to the encoding phase in VR. In the VR version of the distance estimation task, participants reported distances by walking the respective distance between two remembered object positions in a circular enclosure. In the desktop version of this task, they indicated these distances on a subjective scale (Figure 3, see Methods). Participants successfully completed both versions of the task (Supplemental Figure 4AB; long vs short distances in VR version: t(36)=11.00, p=9.999×10^{-5}; mean±SD of Spearman correlations between true and estimated distances in desktop version: r=0.67±0.19), demonstrating the ability to compute

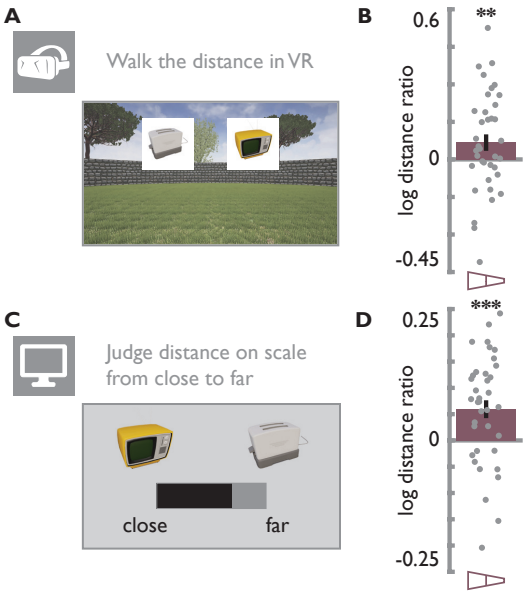


Figure 3 | Distortion of distance estimates.

A. Participants were cued to estimate and walk distances between the remembered positions of object pairs in a circular virtual environment. **B.** Taking advantage of matched distances between object positions, estimated distances were averaged for each triplet and expressed as a distance ratio of estimated distances in the narrow over the broad part of the trapezoid. The log distance ratio differed significantly from 0. **C.** In the desktop version of the task, participants adjusted a slider to estimate the distance between object pairs on a subjective scale from close together to far apart. **D.** The estimates for the same distances differed between the two parts of the trapezoid as indicated by the log distance ratio, which was significantly different from 0. All bars show mean±SEM and grey circles indicate individual subject data. ** $p<0.01$, *** $p<0.001$

never-experienced distances from pairs of remembered positions. Comparing distances walked in the VR version of the task to true Euclidean distances across all trials revealed an overestimation bias ($t(36)=5.78$, $p=9.999\times10^{-5}$).

The regular firing patterns of grid cells are assumed to support vector calculations, allowing navigational vectors to be calculated by comparing the current grid population phase with the phase of remembered locations (Fiete et al., 2008; Erdem and Hasselmo, 2012; Kubie and Fenton, 2012; Bush et al., 2015; Banino et al., 2018). Therefore, the differences in grid cell firing patterns between the narrow and the broad part of the trapezoid (Krupic et al., 2015) might result in differential estimates for distances of the same magnitude. More specifically, differences in the change of grid population phase between equidistant positions might explain such an effect (Carpenter and Barry, 2016) (Figure 1D). Taking advantage of our

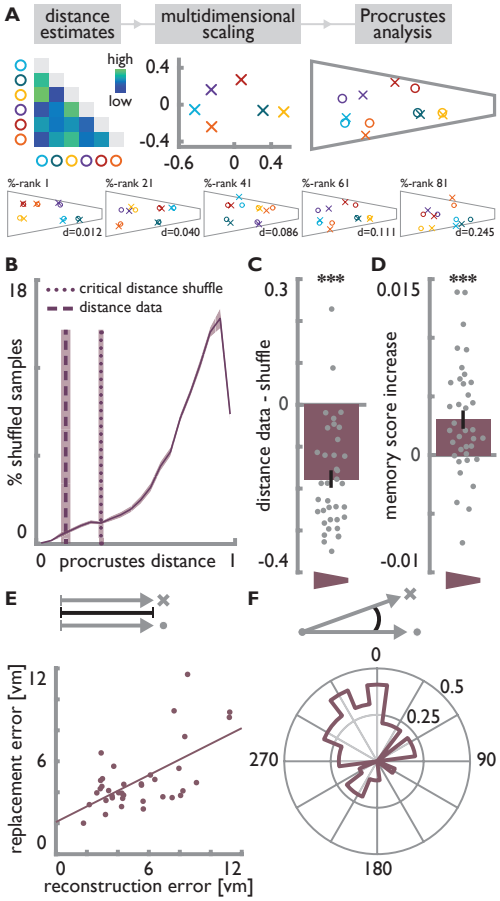


Figure 4 | Reconstruction of remembered positions.

A. Using multidimensional scaling (MDS), coordinates along two underlying dimensions were extracted from pairwise distance estimates. The resulting coordinates were mapped to the original object positions in the trapezoid using Procrustes analysis (see Methods). Color bar indicates estimated distances. Data shown for a randomly selected participant in top panel. Bottom row shows five participants chosen to illustrate reconstruction accuracy across the sample based on the spread of Procrustes distances (data shown for participants at percent ranks 1, 21, 41, 61 and 81 from left to right). Colored circles indicate correct positions and crosses the respective reconstructed positions. **B.** The Procrustes distance quantifies the deviation between true and reconstructed positions as the normalized sum of squared error distances (mean across participants shown by dashed vertical line). For each participant, a surrogate distribution of Procrustes distances was obtained from fitting the coordinates from MDS to coordinate sets with shuffled object identities (solid line). Dotted vertical line indicates the averaged critical Procrustes distance defined as the 5th percentile of the respective surrogate distributions. Shaded areas show SEM across participants. **C.** The Procrustes distances from fitting to true coordinates were significantly smaller than the critical distances of the surrogate distributions. **D.** Memory scores quantifying positional memory within the environment were significantly higher when calculated with respect to the reconstructed rather than the true object positions. **E.** The average lengths of vectors from true object positions to response and reconstructed positions correlate across participants **F.** The angular difference in orientation between vectors from true object positions to response and reconstructed positions clusters around zero. Bars in C and D show mean±SEM and grey circles indicate individual subject data. *** p<0.001

design in which participants learned a triplet of object positions in each half of an environment with matched inter-object distances, we tested this hypothesis by comparing the distances remembered for the two triplets. The log ratio of mean remembered distances between the narrow and broad part of the trapezoid was significantly different from 0 in both the VR (Figure 3B, $t(36)=2.04$, $p=0.001$) and the desktop (Figure 3D, $t(36)=3.49$, $p=9.999\times10^{-5}$) version of the task, indicating that participants consistently estimated the same distances in the two parts of the trapezoid as different. Indeed, the ratio of remembered distances was highly correlated across the two modalities (Supplemental Figure 4C, Spearman $r=0.71$, $p=1.984\times10^{-6}$). In both versions of the task, the distance ratio of the trapezoid differed significantly from the surrogate distributions of distance ratios obtained from the square (Supplemental Figure 4DE, VR: $Z=3.27$, $p=0.001$; desktop: $Z=4.09$, $p=2.138\times10^{-5}$).

Reconstruction of remembered locations

What is the structure of deformed memory maps? To reconstruct remembered object positions from estimated inter-object distances, we applied multidimensional scaling to the data obtained in the desktop version of the task (Figure 4A). We extracted coordinates along two dimensions (see Supplemental Figure 5A) which we mapped onto the true coordinates of the trapezoid using Procrustes analysis to match the two configurations of coordinates (Figure 4A and Supplemental Figure 5B, see Methods). We quantified the deviance between the true and reconstructed positions after Procrustes analysis and compared this Procrustes distance to a surrogate distribution of distances obtained from shuffling object-position-assignments to assess the statistical significance of the reconstruction accuracy (Figure 4B). The observed Procrustes distances were significantly lower than the 5th percentiles of the surrogate distributions (Figure 4C, $t(36)=-8.48$, $p=9.999\times10^{-5}$), reflecting a close match between true and reconstructed positions. Importantly, re-calculating above-described memory scores with the reconstructed positions led to higher scores compared to the true positions (Figure 4D, $t(36)=3.09$, $p=9.999\times10^{-5}$), providing direct evidence that positional memory is used to compute distances between objects and that distorting the spatial map also distorts distance estimates. This effect was also significant when excluding trials targeting objects whose reconstructed position lay outside of the environment ($t(36)=1.42$, $p=0.023$). For each participant, we then calculated the error vectors between the true and remembered positions in the object position memory task and compared these to the error vectors of the reconstructed positions. Overall, there was a strong relationship between reconstructed and remembered positions as indicated by a significant correlation of the average

lengths of the error vectors (Figure 4E, r=0.62, p=4.917×10^{-5}) and a clustering of their directions (Figure 4F, angular difference of vectors significantly clustered around 0: v=13.77, p=0.001).

DISCUSSION

Here, we used immersive VR to demonstrate that environmental geometry can distort human spatial memory. Consistent with environmentally-induced distortions of rodent entorhinal grid-patterns, our data show that positional memory is impaired in a trapezoid compared to a square with the deficits being most pronounced in the narrow end of the trapezoid environment. Importantly, the distortions persisted outside of the environment as participants estimated identical distances to be different between the narrow and broad part of the trapezoid, underscoring the effect of environmental geometry during encoding on subsequent memory. Moreover, remembered positions reconstructed from these distance estimates directly reflected positional memory during the learning task.

Our findings indicate a strong influence of environmental geometry on human spatial memory. This influence has been predicted from rodent data in which the six-fold symmetry of grid firing is distorted in a trapezoidal enclosure, with the most pronounced distortions in the narrow part of the enclosure (Krupic et al., 2015). These irregularities are manifest as changes in the scale and orientation of firing patterns (Krupic et al., 2015). The grid cell population phase is thought to provide a mechanism to encode spatial positions and calculate vectors between locations (Bush et al., 2015). As such, distortions of the grid-pattern will decouple the rate of change in population phase from distance traveled (Carpenter and Barry, 2016). Concurrent with impaired positional decoding from simulated grid cells with locally distorted firing patterns (Krupic et al., 2018), this might lead to uncertainty in the discrimination of positions in space and ultimately result in impaired positional memory and diverging distance estimates for the same Euclidean distances. In concert with evidence for impaired path integration with disrupted grid cell firing in rodents (Gil et al., 2018) and increased path integration errors in older adults with weaker hexadirectional signals measured with fMRI (Stangl et al., 2018), previous studies support the interpretation that the integrity of the grid pattern is beneficial for human spatial memory. The strength of hexadirectional signals and the directional coherence of the orientation of these signals across voxels in the entorhinal cortex are associated with memory performance across participants learning object

positions in circular enclosures (Doeller et al., 2010; Kunz et al., 2015). Our findings dovetail with this notion as they demonstrate that environmental geometry, known to compromise grid-patterns in rodents, influences spatial cognition in a within-subject design.

Prior studies suggested that changing environmental boundaries might influence human spatial cognition in ways consistent with findings from studies of rodent place (O'Keefe and Dostrovsky, 1971) and grid cells (Hafting et al., 2005). Focusing on path integration, one of the core functions assumed for grid cells (Hafting et al., 2005; McNaughton et al., 2006; Moser et al., 2017), biases in human navigation have been reported to follow predictions derived from grid cell firing (Chen et al., 2015). In particular, the experimental design in Chen et al. (2015) built upon the observation that rodent grid-patterns rescale to match changes made to the geometry of already familiar enclosures (Barry et al., 2007). Expansions and compressions of boundaries relative to preceding trials resulted in under- and overshoots of the return path in a path integration task, when the path included a component along the manipulated boundary dimension (Chen et al., 2015). This illustrates how, through environmental change, altering the rate of change in grid cell population phase in relation to distance traveled can introduce biases in human navigation (Carpenter and Barry, 2016; Chen et al., 2015). Translating this idea to the memory-based estimation of distances between locations might explain the diverging judgments of identical distances observed in our data. Expansions and compressions of virtual environments have further been demonstrated to impact spatial memory in humans and under conditions of environmental change, positional memory follows models of place cells and boundary proximity (Hartley et al., 2004; Schuck et al., 2015a). While the studies described above indicate how boundary manipulations in familiar environments influence spatial behavior, we built upon work showing that distorted grid-patterns persist in static trapezoid environments even with prolonged experience (Krupic et al., 2015). Our findings suggest that distortions of the brain's spatial metric can result in mnemonic distortions under constant boundary conditions within a specific environment and even outside of this encoding environment.

We opted for a purely behavioral experiment; our hypotheses, experimental design and analysis however directly built upon findings from electrophysiological recordings of grid cells in rodents (Krupic et al., 2015). We employed highly immersive VR technology to enhance the impact of environmental geometry on spatial cognition and engage proprioceptive, vestibular and motor systems during the task. Currently, immersive VR does not allow the concurrent recording of neural

data. The contribution of locomotor cues to the experience of navigation in general has been emphasized previously (Taube et al., 2013) and a recent study in rodents has used gain manipulations in VR to emphasize the contributions of locomotor cues to grid cell firing specifically (Campbell et al., 2018). Having established the impact of environmental geometry on human spatial cognition, an exciting question for future research would be to combine manipulations of environmental geometry with neuroimaging techniques such as fMRI to study the deformations of the cognitive map we describe here in the brain. To do so, an important measure could be the hexadirectional signal that can be observed in the human entorhinal cortex (Doeller et al., 2010). Beyond fMRI, an exciting future avenue is paved by the development of new magnetoencephalography systems, which might allow the combination of immersive VR with recordings of neural data (Boto et al., 2018).

As large parts of human indoor navigation take place in rectangular rooms, the novelty of a trapezoidal enclosure in our task might be considered as a factor contributing to impaired performance compared to the square. Such an effect of unfamiliarity with polarized environments, however, would not predict the observed within-environment differences in performance. Further, participants' walking speeds did not differ between environments and their paths from the start to the remembered object position were not more or less direct. Thus, none of our control measures provided evidence for fundamental differences in navigational performance between the environments per se. The absence of differences in navigation performance is in line with the successful use of trapezoidal room geometry during spatial updating (Kelly et al., 2008). Additionally, the detection and encoding of color change events was not affected by the environmental manipulation, speaking against an effect of increased task demand in the trapezoid as sufficient attentional resources were available for this secondary task.

Importantly, the effects we observed in positional memory persisted outside of the environment as demonstrated by the differential estimates for matched distances between positions within the different parts of the trapezoid. These distortions were consistent across response modalities. The response profiles observed in the VR version of the task revealed a general tendency to overestimate distances between positions, in line with previous studies reporting overestimations of navigated distances (Brunec et al., 2017) and spatial scale in map drawings (Jafarpour and Spiers, 2017). We used the distances estimated on a subjective scale in the desktop version of the task to reconstruct remembered positions. Accounting for the distortions in participants' memory by using these reconstructed positions to re-compute memory

scores yielded increased performance scores. This illustrates the close match between positions reconstructed from distance estimates and positional memory within the environment, and demonstrates that, consistent with the formation of cognitive maps (O'Keefe and Nadel, 1978), distances never directly experienced in the task were computed from remembered positions. Grid cells have been suggested to support this kind of vector computation (Bush et al., 2015; Banino et al., 2018). This is further in line with evidence for the involvement of the entorhinal grid system in imagination (Bellmund et al., 2016; Horner et al., 2016) and theoretical accounts proposing a role for spatially tuned cells in memory (Byrne et al., 2007; Buckner, 2010; Hasselmo, 2011).

The persistent effects of environmental geometry outside of the highly polarized environment also speak to a growing body of literature implicating grid cells in mapping continuous, task-relevant dimensions beyond navigable space. Grid cells have been implicated in mapping visual space in monkeys (Killian et al., 2012) and humans (Nau et al., 2018a; Julian et al., 2018) as well as time during running in place throughout a delay (Kraus et al., 2015) and tone frequency during sound sweeps towards a target frequency (Aronov et al., 2017) in rodents. In humans, hexadirectional signals were also observed during trajectories through an abstract space spanned by the dimensions of neck and leg length of stick figure birds (Constantinescu et al., 2016). Collectively, these findings point towards a role of the entorhinal grid system in representing dimensions of task feature spaces. As proposed for navigable space (Hafting et al., 2005; McNaughton et al., 2006; Bush et al., 2015; Moser et al., 2017; Herz et al., 2017), the regular firing patterns of grid cells might provide a metric for these spaces allowing the efficient encoding of specific stimuli located at different positions within a space. Speculatively, correlated feature dimensions or feature spaces in which subsets of feature combinations are impossible might distort how grid cells map these spaces in a similar way as environmental geometry distorts grid cell firing patterns, resulting in biased representations similar to the distortions of spatial memory observed in this study.

In conclusion, our data show distortions of human spatial memory consistent with the changes induced in rodent grid cell activity by the geometry of highly polarized enclosures. These distortions persist outside of the environment, indicating an enduring impact of environmental geometry on memory. In line with the proposed roles for grid cells in navigation and mapping feature dimensions beyond navigable space, these findings suggest that environmental geometry might be able to distort the metric of cognitive representations.

METHODS

Participants

53 Participants between the age of 18 and 30 were recruited from the Norwegian University of Science and Technology. All participants provided written informed consent before participation, and all research procedures were approved by the regional ethics committee (REC North, reference number 2017/153). Sample size was based on a power calculation assuming a small to medium effect (d=0.4) of environmental geometry on human spatial cognition, resulting in a sample size of 52 to achieve a statistical power of 80% ($\propto$=0.05, two-tailed test). 39 participants (mean age 23.8±2.5 years, 36% female) completed the experiment (14 incomplete datasets due to technical difficulties with the VR setup or motion sickness). Two participants were excluded due to poor memory performance defined as average replacement errors more than 1.5 times the interquartile range larger than the upper quartile of average errors in the sample. Thus, 37 participants entered the analyses.

Overview

Over the course of the experiment, participants completed the object position memory task in two environments and estimated durations between (color change) events occurring during these navigation tasks as well as distances between pairs of object positions. The general structure was as follows: Participants were first familiarized with the VR setup and the time estimation task before beginning the object position memory task in the first environment. The object position memory task was carried out in a trapezoidal or square environment for 20 minutes each, with the order of environments counterbalanced across participants. Subsequent to navigating an environment, participants were prompted to estimate the durations between color change events encountered in that environment. In the last two tasks, participants were prompted to estimate distances between pairs of objects in VR and on a computer screen, respectively. The design of each task and the corresponding analyses are described in detail in the following sections. All analyses were performed using Matlab (Release 2017a, The MathWorks, Inc.) and statistical tests (two-tailed unless stated otherwise) were performed using the Resampling Statistical Toolkit (http://www.mathworks.com/matlabcentral/fileexchange/27960-resampling-statistical-toolkit). Specifically, test statistics were compared against a surrogate distribution obtained from 10000 bootstrap samples respecting within-subject dependencies. Circular statistics were implemented using the Matlab-based Circular Statistics Toolbox (Berens, 2009).

Virtual reality

Aiming to maximize the feeling of immersion and thereby the impact of environmental features we employed state of the art VR technology consisting of a head mounted display (HMD, Oculus Rift CV1) and a motion platform (Cyberith Virtualizer). Participants wore low-friction overshoes and were strapped into a harness attached to the motion platform's ring system allowing free rotations. To navigate the virtual environments, participants were instructed to lean slightly into the ring construction to slide the front foot backwards across the sensors of the low-friction base plate of the motion platform while taking a step forward with the back foot (see Supplementary Video 1), generating translational movement in the current forward direction determined by the orientation of the participant in the ring system (Cakmak and Hager, 2014). Head movements were tracked in 3D using the HMD's tracking system and the virtual environments were displayed to both eyes separately at a resolution of 1080 x 1200 pixels and a refresh rate of 90 Hz. The virtual environments were created and presented using the Unreal Engine (v.4.13.2, Epic Games Inc., 2017) and participants' eye height was set to 1.80 virtual meters (vm). Participants were familiarized with the VR setup in a circular environment (45.74vm in diameter) consisting of a grass floor curtailed by a wall (height 3.75vm). A set of trees spread around the outside the environment served as cues for orientation. During familiarization, participants practiced walking and turning by navigating the circular environment to collect coins appearing at random positions in the environment. Participants were instructed to walk towards the coins and collect them via button presses on a handheld controller. Additionally, this familiarization period served as a practice for the time estimation task (see below).

Object position memory task

Participants performed an object position memory task during which they iteratively learned the positions of six objects in a trapezoidal environment (36vm×76vm×8vm×76vm) with side lengths proportional to the enclosure rodents explored in a study reporting distortions of grid cell firing patterns (Krupic et al., 2015). To establish a behavioral baseline, participants performed this task also in a square control environment (40.27vm×40.27vm) with equal surface area. There were no distal cues outside of the environment to enforce spatial learning based on environmental geometry. To facilitate orientation, each wall was presented in a unique color. Both environments had a grass floor and a blue sky with moving clouds was visible (Figure 1A). Participants performed the task for 20 minutes in each environment with the order counterbalanced across participants. In each environment, participants learned the positions of six everyday objects presented

as three-dimensional models. The assignment of objects to arenas and positions was randomized across participants.

In each trial of an initial learning phase, participants navigated to a start position indicated by a traffic cone. Then, an object was shown at its predefined position in the environment and participants were instructed to navigate to the object, collect it via button press and memorize its position. Each object was shown once and the order of objects was randomized. In the subsequent test phase (Figure 1A), participants again navigated to start positions. Upon arrival, a picture of one of the objects was shown as a cue for 3 seconds in front of the participant, prompting participants to navigate to where they remembered this object in the environment. Participants indicated the remembered position via button press after arrival and received feedback about their accuracy in the form of one of five smiley faces. The object then appeared at its correct position and participants collected it before the beginning of the next trial. Participants completed 30.54±6.71 and 30.38±8.09 (mean±SD) test trials in square and trapezoid, respectively (number of trials not significantly different: $t(36)=0.18$, $p=0.759$).

The order of trials was randomized for mini-blocks of six trials, so that within a mini-block each object was sampled once and no two consecutive trials sampled the same objects. A triplet of object positions (Figure 1B) was randomly generated for each participant with a minimum distance of 11vm between object positions and a minimum of at least 3vm to the nearest boundary. Positions were constrained so that the connection between two objects was parallel to the long-axis of the trapezoid or one of the walls of the square. The third object was placed at an angle ranging from 90°-120° relative to the first two with the same distance to one of the objects as between the first two. Such a triplet of positions was placed in both the narrow and broad part of the trapezoid defined based on the midpoint of its long-axis and the left and right part of the square. Placing triplets of objects with matched distances in each part of the environment allowed direct comparisons of remembered distances between environments and their sub-parts (see distance estimation tasks). Since cues were only shown once participants arrived at the start position of a given trial, participants never walked the direct path between two objects. Distances from start to target object positions (mean and standard deviation square: 18.66±4.65vm; trapezoid: 19.92±8.50vm; trapezoid broad: 21.10±10.95vm; trapezoid narrow: 18.73±4.67vm) did not influence spatial memory performance (Supplemental Figure 2EF).

Positional memory

Replacement errors were quantified as the Euclidean distance between the correct position of an object in the environment and the position remembered by the participant. To limit the influence of outlier trials we excluded trials with replacement errors larger than 1.5 times the interquartile distance above the upper quartile of errors for each participant (mean±SEM number of trials excluded =3.35±0.26) from all further analyses. Average replacement errors were compared across environments using a bootstrap-based paired t-test (Figure 2B). To account for the fact that despite equal area larger replacement errors are possible in the trapezoid compared to the square control environment, we subsequently quantified performance using memory scores. Specifically, we generated a distribution of 1000 random locations uniformly covering each environment and quantified for each trial the proportion of locations further away from the correct object position than the position indicated by the participant. Importantly, calculating memory scores based on the distribution of possible errors for each target position yields a measure comparable across positions and environments (Jacobs et al., 2016) with a chance level of 0.5 for random performance and scores closer to 1 for high performance. To test the hypothesis of degraded spatial memory in the trapezoid memory scores were compared across environments using a bootstrap-based paired t-test (Figure 2C).

In a next step, we aimed to test the more specific hypothesis of increased degradation of positional memory in the narrow compared to the broad part of the trapezoid derived from the larger distortions of firing patterns of grid cells in this part of the environment (Krupic et al., 2015). To account for the high performance levels of our participants compared to the baseline of chance performance and differences in local geometry, we referenced performance in the trapezoid to performance in the square control environment for this analysis (Figure 2D). Specifically, the 99th percentile of replacement errors across all participants and trials served as a reference value to quantify performance in each single trial as 1 minus the quotient of a trial's error and the reference value, resulting in higher scores for better performance. Trials with errors larger than the reference value were assigned a score of 0, attenuating the effects of trials with very large errors. For each participant, we calculated the ratio of scores averaged across object positions in the narrow and broad part of the trapezoid, respectively. After log transformation, we again used a bootstrap-based t-test to test whether the resulting ratio was significantly below 0 (Figure 2E). Outlier participants were excluded based on our standard criterion of values more than 1.5 times the interquartile range above or below the upper or lower quartile, respectively (see Supplemental Figure 1B for full dataset). The memory ratio also

differed significantly from 0 when calculated using raw errors (Supplemental Figure 1B). Note that when calculating the memory ratio based on raw error scores we divided errors in the broad by errors in the narrow part of the trapezoid so that low values of the memory ratio consistently reflected impaired performance in the narrow part of the environment. Since the rotationally symmetric geometry of the square does not pre-define how to calculate the memory ratio, we created a surrogate distribution by shuffling which half of the environment was to serve as the numerator and denominator for the memory ratio across participants. For each permutation, we calculated the ratio between memory performance (expressed relative to the same reference value as for the trapezoid) for objects located in the two halves of the square. The log memory ratio observed in the trapezoid was smaller than the 5th percentile (one-tailed test) of the surrogate distribution obtained from 10000 permutations, (Supplemental Figure 1C). The shape of the surrogate distribution did not differ from normality (Kolmogorov-Smirnov test, D=0.001, p=0.521), we hence used it to convert the p-value reflecting the number of occurrences of smaller memory ratios in the surrogate distribution into a Z-statistic. To visualize response behavior in the two parts of the trapezoid we collapsed across all trials from all participants for objects located in the broad and narrow part of the arena. Response positions were centered on the respective true positions and divided into 50x50 square bins with a side length of 0.6vm. The resulting histogram was smoothed using a Gaussian kernel with a standard deviation of 0.5vm and plotted as a heatmap (Supplemental Figure 1D).

Parameters of navigation

To assess whether differences in navigation behavior might underlie the observed differences in positional memory, we analyzed navigational performance in the replacement phase of each trial, where participants navigated to the remembered position of a cued object. For each trial, we calculated the Euclidean distance between the start position and the response location and subtracted it from the length of the path walked by the participant. This excess path length measures the directness of the paths taken, potentially reflecting the degree of certainty about the trajectory as increased uncertainty might lead to more turns and longer paths. We contrasted averaged excess path lengths between the two environments and the broad and narrow part of the trapezoid (Supplemental Figure 2AB). Likewise, we contrasted average walking speeds during the replacement phase between the environments and trials targeting objects from the two trapezoid parts (Supplemental Figure 2CD).

Further, we assessed whether the distance from a trial's start position was related to the accuracy of object position memory in a consistent way across subjects. For each subject, we calculated the Spearman correlation coefficient between the distances from start to true object positions and positional memory as defined by the Euclidean distances between true and remembered object positions. The resulting coefficients were tested against 0 for all trials in the two environments separately (Supplemental Figure 2E) or for trials probing objects in the narrow and broad part of the trapezoid, respectively (Supplemental Figure 2F).

In a next step, we assessed rotations participants made during the replacement phase of the trial. To this end, we centered the rotation of the body as measured by the orientation of the motion platform's ring construction and the orientation of the participant's head as tracked by the HMD on the direction from start to response position. We averaged orientation values for trials within square and trapezoid or broad and narrow part of the trapezoid, respectively, and tested for clustering around 0 using V-tests and differences of averaged orientation values between conditions using Watson-Williams tests (Berens, 2009) (Supplemental Figure 3, top row). Additionally, we quantified the circular variance of centered orientation values and contrasted it across conditions (Supplemental Figure 3, bottom row). None of these measures suggested influences of navigation behavior per se on the key conclusions of the paper.

Distance estimation tasks

After completing the time estimation task following the second object position memory task in the second environment, participants estimated distances between pairs of object positions in two modalities: on a computer screen and by walking the actual distances in VR.

Virtual reality

Participants were placed in the same circular virtual arena as during the familiarization session. Each trial began with an arrow pointing to the middle of the arena, with the arrow appearing at a random location on the arena floor. After participants positioned themselves on the base of the arrow, images of two objects were presented in front of them for 3 seconds (Figure 3A). Participants were instructed to walk the distance they remembered the objects to be apart based on the object position memory task while following the direction indicated by the arrow. When participants terminated a trial via button press, a checkmark was presented to indicate the successful registration of the response and the next trial began. Due to

time constraints this task was restricted to distances between objects within a triplet, resulting in 12 trials making up a block. Trial order within blocks was randomized with the constraint that trials with objects from the two environments alternated. Participants completed two blocks with a short break in between.

Since only distances within a triplet of positions were tested in this task, participants' averaged estimates for the long and short distances were compared using a bootstrap-based paired t-test as an indicator of successful task performance (Supplemental Figure 4A). To test whether distance estimates for the same distances differed between narrow and broad part of the trapezoid, we took advantage of the fact that true distances were matched across position triplets and thereby arena parts. Response distances within a triplet were averaged and expressed as the log ratio of mean distances in narrow over broad part for statistical comparisons using a bootstrap-based t-test against 0 (Figure 3B). As for the memory ratio, we created a surrogate distribution to compare the distance ratio observed in the trapezoid to the square by shuffling across participants which half of the square was to serve as the numerator and denominator for the distance ratio in each of 10000 permutations. The log distance ratio observed in the trapezoid was more extreme than the 2.5th and 97.5th percentiles (two-tailed test) of this surrogate distribution (Supplemental Figure 4D). The shape of the surrogate distribution did not differ from normality (Kolmogorov-Smirnov test, D=0.01, p=0.287).

Computer monitor

Afterwards, participants were instructed to estimate distances between object pairs on a desktop computer setup. Images of objects on a white background, as well as an adjustable horizontal bar with the labels 'close together' on the left and 'far apart' on the right were presented on a computer screen (Figure 3C). Again, participants were instructed to estimate how far objects were apart during the object location memory task. Here, they indicated their response by adjusting the horizontal bar with a computer mouse, after which a grey screen was shown for 500 milliseconds. All possible combinations of distances were probed, i.e. also comparisons across triplets, yielding subjective distances between all pairs of object positions in an environment. Each of the 15 combinations of object pairs per environment was probed twice, resulting in a total of 60 trials. Trial order was randomized with the constraint that each possible pair of objects was sampled before any object combination was sampled for the second time. Each object was shown once on the left and once on the right side of the screen in the two trials sampling a given object pair. This distance estimation task as well as the time estimation task was presented

using the Psychophysics Toolbox (Brainard, 1997) for Matlab (Release 2016a, The MathWorks, Inc.). General performance in this task was assessed by calculating Spearman correlations between the estimated distances and the respective true distances (Supplemental Figure 4B). Further, the remembered distance ratio was calculated and tested in the same way as described above. The surrogate distribution obtained for comparison to the square did not differ from normality (Kolmogorov-Smirnov test, D=0.01, p=0.186).

Reconstructing remembered positions

To reconstruct remembered object positions in the trapezoid from distance estimates, multidimensional scaling (MDS) was applied to the distance estimates obtained in the desktop version of the task as only here distances between all pairs of positions were estimated. Estimated distances were normalized to a range from 0 to 1 and averaged across the two repetitions of each object pair and subjected to MDS to recover coordinates reflecting this distance structure using metric stress as the cost function and a random initial configuration of points. Our approach assumes that two dimensions underlie the object location memory formed during the navigation task. To assess whether this assumption holds, we compared the model deviance of general linear models predicting the distances between true positions from the positions recovered from MDS for different numbers of dimensions. As expected, unexplained variance was substantially decreased when using two instead of one dimension, but no clear improvement resulted from a larger number of dimensions (Supplemental Figure 5).

To match the coordinates resulting from MDS to the original positions in the virtual environment we used Procrustes analysis allowing translation, scaling, reflection and rotation. The goodness of fit, the Procrustes distance, was quantified by the normalized sum of squared errors between reconstructed and true coordinates and was compared to Procrustes distances resulting from Procrustes analyses of the MDS coordinates and sets of coordinates in which the assignment of object identity to position was shuffled, yielding a surrogate distribution from all 720 possible permutations. Specifically, we tested on the group level whether the fits between reconstructed coordinates and true coordinates were better than the fits constituting the 5th percentile (reflecting the threshold for statistical significance at an alpha-level of 0.05) of each participant's surrogate distribution (Figure 4BC). The reconstructed coordinates are visualized as heatmaps in Supplemental Figure 5B following the same procedure as described above.

To test whether the reconstructed positions indeed reflected participants' memory in the object position memory task, we re-calculated the memory scores as described above but with the coordinates resulting from the Procrustes analysis instead of the true object positions as goal positions (Figure 4D). To rule out that objects whose positions were reconstructed to be remembered outside of the environment were driving the effect, we excluded all affected trials from the memory score calculation in an additional control analysis. To describe the overlap between positions reconstructed from distance estimates and performance in the object position memory task, we calculated error vectors based on the true object positions and quantified the match between average error vectors of response and reconstructed positions by correlating their lengths using Pearson correlation (Figure 4E). We further probed these error vectors' similarity in orientation by averaging the angular differences between vectors from the correct to the respective response and reconstructed positions for each participant and testing the resulting circular means for a clustering around 0 using a V-test (Figure 4F).

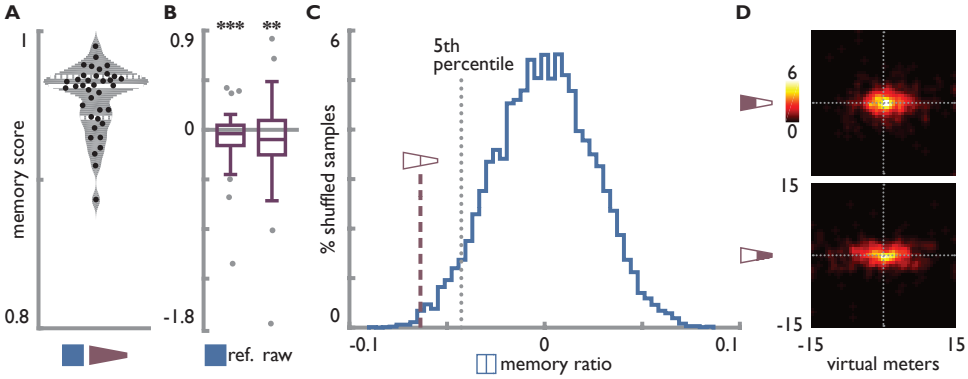
Time estimation task

Here, we asked participants to perform a secondary task during the navigation period in each arena. In the sky above the arena a ring was presented, which changed color four times during the object position memory task per environment. The ring remained in a given color for an interval between 2 and 6 minutes and participants indicated color changes via button presses and were instructed to remember the order of colors and the duration for which each color was presented. While different colors were presented in the two environments, the intervals between color changes were constant across environments allowing for a comparison of temporal memory between square and trapezoid.

After completing the object position memory task in the first environment, participants were placed in front of a computer screen to estimate the time between color changes before continuing the object position memory task. On a white screen, two pairs of consecutive colors were shown and participants indicated the time interval they remembered to separate the two color changes in minutes and seconds, e.g. how much time passed between the ring changing color from blue to yellow and changing from yellow to green. Participants were cued to estimate the time between all six possible combinations of color changes per environment. To ensure full understanding of this task, participants estimated intervals between color changes occurring at random times between every 30 and 120 seconds during the familiarization phase prior to the object position memory task. Overall performance

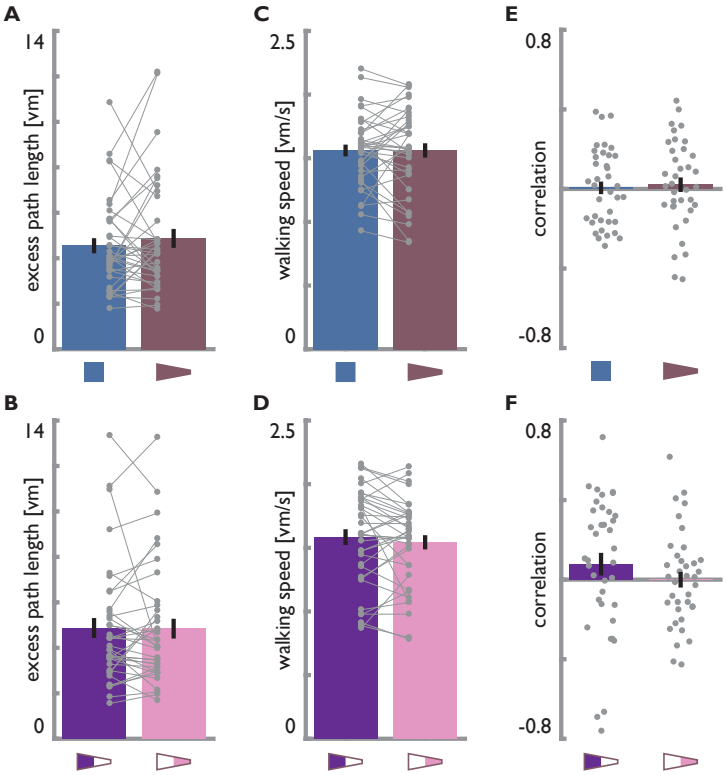
in this task was quantified using Spearman correlations between the correct and estimated time intervals before specifically comparing average estimation errors, absolute estimation errors and the standard deviation of estimation errors across environments (Supplemental Figure 6).

SUPPLEMENTAL FIGURES



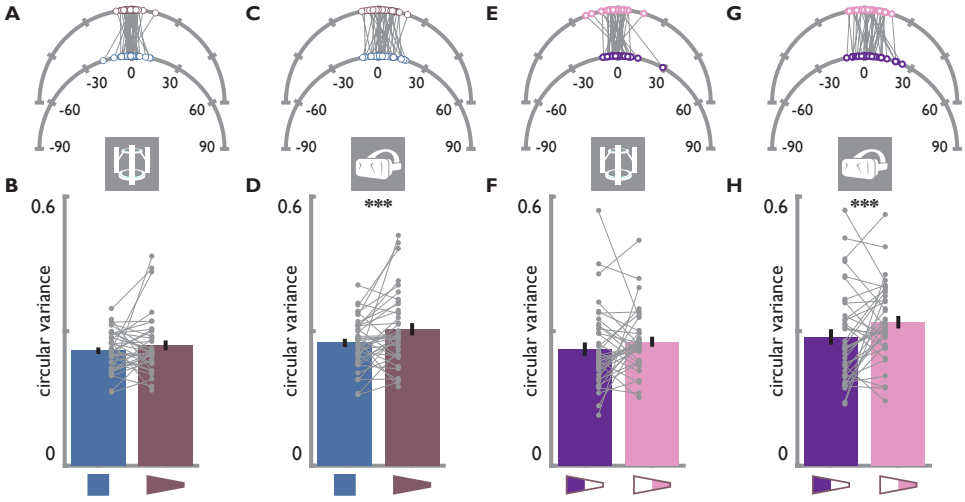
Supplemental Figure 1 | Positional memory.

A. Distribution of average memory scores across participants. Grey area indicates normal kernel density estimate, solid white line shows median and dashed white lines show upper and lower quartile of distribution. Black circles show memory scores of individual participants. **B.** Log memory ratio for the two parts of the trapezoid for performance referenced to square behavior (left) and for raw error scores (right). Lower values indicate lower performance in the narrow part of the trapezoid and log memory ratios are consistently below zero (referenced to square: $t(30)=-2.52$, $p=9.999\times10^{-5}$; raw error ratio: $t(32)=-1.89$, $p=0.006$). Data points more than 1.5 times the interquartile range above or below the upper or lower quartile were excluded as outliers (grey dots), but comparable results are obtained without outlier exclusion (referenced to square: $t(36)=-1.85$, $p=0.005$; raw error ratio: $t(36)=-1.44$, $p=0.032$). Boxplots represent median as well as upper and lower quartile of distribution, whiskers show most extreme value within 1.5 times the interquartile range from the upper and lower quartile respectively. **C.** The log memory ratio observed between the trapezoid parts (dashed line represents mean of ratio scores shown in Fig. 2E) was significantly lower than the critical value (5th percentile, dotted line) of a shuffle distribution (blue) obtained from computing the log memory ratio between the square halves across 10000 iterations. **D.** Heatmaps showing response locations for all trials across all participants for objects in the broad (top) and narrow (bottom) part of the trapezoid. Dotted lines show correct location in x- and y-dimension with their intersection representing the true position. ** $p<0.01$, *** $p<0.001$



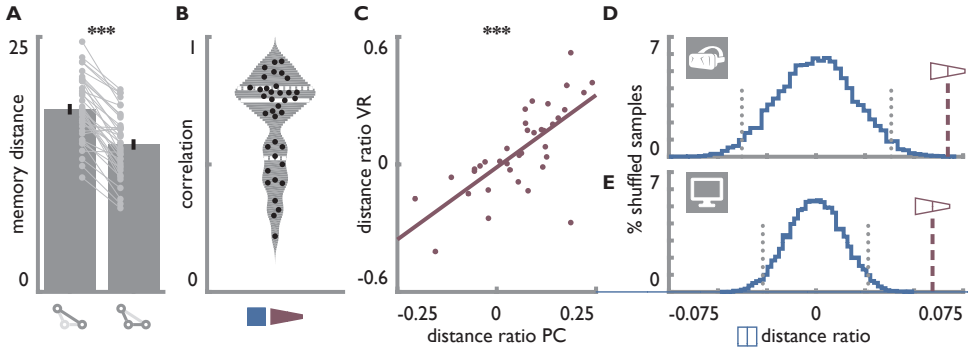
Supplemental Figure 2 | Navigational variables do not differ between environments.

A,B. The excess path length of the trajectory from start to response position did not differ between square and trapezoid or the two parts of the trapezoid. **C,D.** Walking speed did not differ between square and trapezoid or the two parts of the trapezoid. **E,F.** Spearman correlation coefficients between the Euclidean distance from the start to the correct object positions and replacement error do not differ from 0 across subjects in the square or trapezoid or for objects located in the broad and narrow part of the trapezoid separately. Bars show mean±SEM and grey circles indicate individual subject data with lines connecting data points from the same participant.



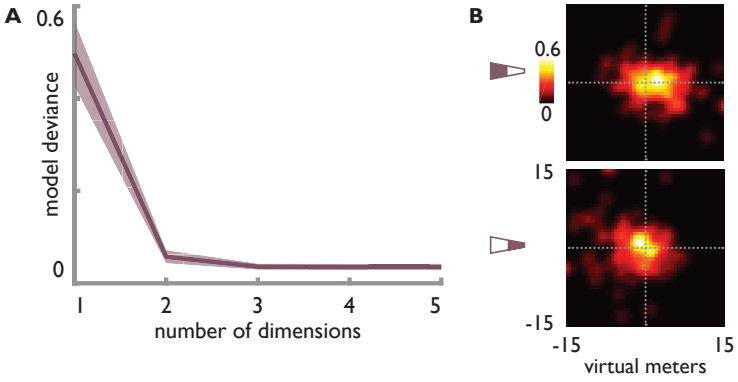
Supplemental Figure 3 | Head and body orientation during navigation.

A. Circular means of body rotations centered on each trial's direction from start to response position. Means are significantly clustered around 0 for both square and trapezoid and do not differ significantly. **B.** Circular variance of body rotations over trials averaged for each participant does not differ between square and trapezoid. **C.** Circular means of head rotations centered on each trial's direction from start to response position. Means are significantly clustered around 0 for both square and trapezoid and do not differ significantly. **D.** Circular variance of head rotations over trials averaged for each participant is larger in the trapezoid than in the square. **E.** Circular means of body rotations centered on each trial's direction from start to response position. Means are significantly clustered around 0 for trials with target object positions in the broad and narrow part of the trapezoid, respectively, and do not differ significantly. **F.** Circular variance of body rotations over trials averaged for each participant does not differ between navigation periods for target objects located in the broad or narrow portion of the trapezoid. **G.** Circular means of head rotations centered on each trial's direction from start to response position. Means are significantly clustered around 0 for true positions in the broad and narrow part of the trapezoid and do not differ significantly. **H.** Circular variance of head rotations over trials averaged for each participant is smaller when cued object position is in the broad compared to the narrow portion of the trapezoid. *** $p<0.001$



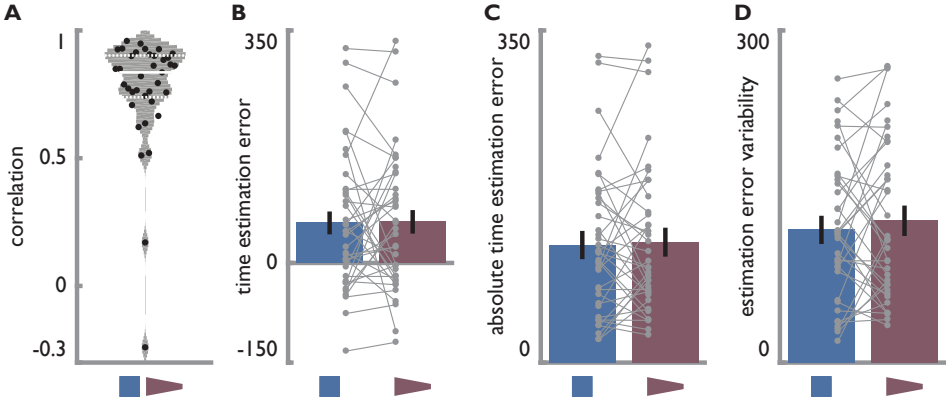
Supplemental Figure 4 | Distance estimates.

A. Long distances (i.e. the base of the isosceles triangle formed by a triplet of positions) were estimated to be longer than the shorter distances (i.e. the legs of the isosceles triangle). Only within-triplet distances were estimated in VR. Bars show mean±SEM and grey circles indicate individual subject data with lines connecting data points from the same participant. **B.** Grey area indicates distribution of Spearman correlation (mean±SD r=0.69±0.19) coefficients between correct and estimated distances based on normal kernel density estimate. Solid white line shows median and dashed white lines show upper and lower quartile. Black circles show correlation coefficients of individual participants. **C.** Significant Spearman correlation of log distance ratio comparing the two parts of the trapezoid obtained from distance estimates on the desktop and in VR. Circles denote individual participant data. **D,E.** The log distance ratio observed between the trapezoid parts (dashed line) was more extreme than the critical values (dotted line) of the shuffle distribution (blue) obtained from computing the log distance ratio between the square halves across 10000 iterations for the distance estimates in VR (**D**) and on the PC (**E**).



Supplemental Figure 5 | Two dimensions underlie distance estimates.

A. Model deviance of GLMs using pairwise Euclidean distances of coordinates obtained from MDS to predict estimated distances for different numbers of dimensions (solid line shows mean model deviance across participants, shaded area indicates SEM). In line with our a priori assumption that two dimensions underlie the distance estimates, model deviance sharply drops when using two rather than one dimension and there is no substantial benefit from including three or more dimensions. **B.** Heatmaps showing positions reconstructed using multi-dimensional scaling and Procrustes transform for objects in the broad (top) and narrow (bottom) part of the trapezoid. Dotted lines show correct position in x- and y-dimension with their intersection representing the true position.

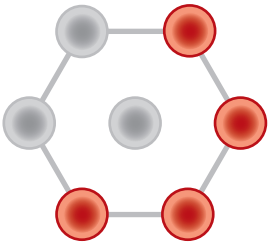


Supplemental Figure 6 | Time estimates do not differ between environments.

A. Grey area indicates distribution of Spearman correlation coefficients (mean±SD r=0.77±0.23) between true and estimated times based on normal kernel density estimate. Solid white line shows median and dashed white lines show upper and lower quartile. Black circles show correlation coefficients of individual participants. **B.** Averaged time estimation errors did not differ between the two environments. **C.** Averaged absolute time estimation errors did not differ between the two environments. **D.** The variability of time estimates as measured by their standard deviation did not differ between environments. Bars show mean±SEM and grey circles indicate individual subject data with lines connecting data points from the same participant.

CHAPTER 4

Structuring Time in Human Lateral Entorhinal Cortex



This chapter is accepted for publication:

J. L. S. Bellmund, L. Deuker, C. F. Doeller (in press). Mapping sequence structure in the human lateral entorhinal cortex. *eLife*. Preprint available at https://doi.org/10.1101/458133.

ABSTRACT

Remembering event sequences is central to episodic memory and thought to be supported by the hippocampal-entorhinal region. We previously demonstrated that the hippocampus maps spatial and temporal distances between events encountered along a fixed route through a virtual city (Deuker et al., 2016), but the content of entorhinal mnemonic representations remains unclear. Here, we demonstrate that, after learning, multi-voxel representations in the anterior-lateral entorhinal cortex (alEC) specifically reflect the temporal event structure. Holistic representations of the temporal structure related to memory recall and the temporal event structure could be reconstructed from entorhinal multi-voxel patterns. Our findings demonstrate representations of temporal structure in the alEC in line with temporal information carried by population signals in the lateral entorhinal cortex of navigating rodents and activations of its human homologue during temporal memory retrieval. Our results provide novel evidence for the role of the human alEC in representing time for episodic memory.

INTRODUCTION

Knowledge of the temporal structure of events is central to our experience. We remember the sequence of events and can recall when in time events occurred. Emphasizing both when in time and where in space events came to pass, episodic memories typically comprise event information linked to a spatiotemporal context. Space and time have been suggested to constitute fundamental dimensions along which our experience unfolds (Konkel and Cohen, 2009; Ekstrom and Ranganath, 2017; Bellmund et al., 2018a). Consistently, the role of the hippocampus — a core structure for episodic memory (Scoville and Milner, 1957; Squire, 1982) — in coding locations in space (O'Keefe and Dostrovsky, 1971; Moser et al., 2017; Epstein et al., 2017) and moments in time (Pastalkova et al., 2008; MacDonald et al., 2011; Eichenbaum, 2014; Ranganath, 2018; Howard, 2018) is well-established. Human memory research has highlighted the role of the hippocampus in the encoding, representation and retrieval of temporal relations that comprise a context (Tubridy and Davachi, 2011; DuBrow and Davachi, 2014; Ezzyat and Davachi, 2014; Hsieh et al., 2014; Jenkins and Ranganath, 2010, 2016; Kyle et al., 2015; Copara et al., 2014). The similarity patterns of mnemonic representations suggest that the hippocampus forms integrated maps reflecting the temporal and spatial structure of event memories (Deuker et al., 2016; Nielson et al., 2015).

How do representations of temporal structure arise in the hippocampus? Evidence suggests that neural ensembles in the lateral entorhinal cortex (EC), which is strongly connected to the hippocampus (Witter et al., 2017b), carry temporal information in freely moving rodents (Tsao et al., 2018). Specifically, temporal information could be decoded from population activity with high accuracy (Tsao et al., 2018). Recently, the human anterior-lateral entorhinal cortex (alEC), the homologue region of the rodent lateral entorhinal cortex (Navarro Schröder et al., 2015; Maass et al., 2015), as well as the perirhinal cortex and a network of brain regions including the hippocampus, the medial prefrontal cortex, posterior cingulate cortex and angular gyrus have been implicated in the recall of temporal information (Montchal et al., 2019). These regions responded more strongly for high compared to low accuracy retrieval of when in time snapshots from a sitcom appeared over the course of the episode viewed in the experiment (Montchal et al., 2019). Together, these findings demonstrate that entorhinal population activity carries temporal information in rodents and that its human homologue is activated during temporal memory recall, but the contents of mnemonic representations in the alEC remains unclear.

To address the question how mnemonic representations in the alEC are shaped by learning a temporal event structure we use representational similarity analysis of fMRI multi-voxel patterns in the entorhinal cortex. Using this paradigm and data, we previously demonstrated that participants can successfully recall the subjective, remembered spatial and temporal relations between object pairs and that the change of hippocampal representations reflects an integrated event map of the remembered distance structure (Deuker et al., 2016). Here, we show that the change of multi-voxel pattern similarity through learning in the alEC specifically reflects the temporal structure of the task.

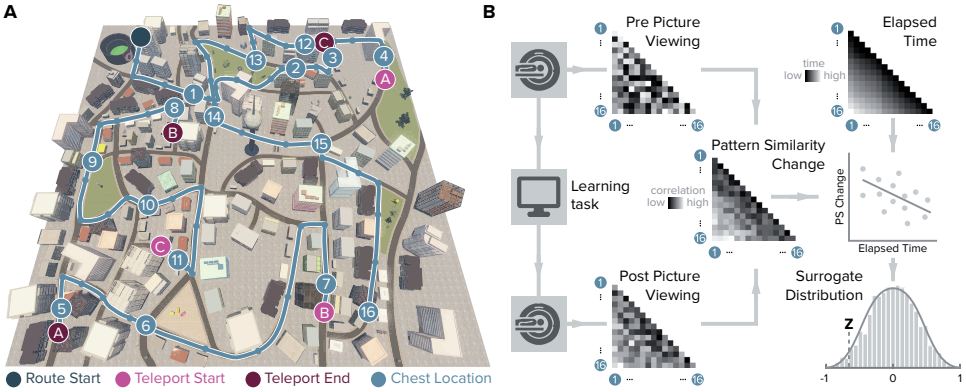


Figure 1 | Design and analysis logic.

A. During the spatio-temporal learning task, which took place in between two identical runs of a picture viewing task (Figure 1—figure supplement 1), participants repeatedly navigated a fixed route (blue line, mean ± standard deviation of median time per lap 264.6±47.8s) through the virtual city along which they encountered objects hidden in chests (numbered circles) (Deuker et al., 2016). Temporal (median time elapsed) and spatial (Euclidean) distances between objects were dissociated through the use of three teleporters (lettered circles) along the route (Figure 1—figure supplement 2), which instantaneously changed the participant's location to a different part of the city. **B.** In the picture viewing tasks, participants viewed randomly ordered images of the objects encountered along the route while fMRI data were acquired. We quantified multi-voxel pattern similarity change between pairwise object comparisons from before to after learning the temporal and spatial relationships between objects in subregions of the entorhinal cortex. We tested whether pattern similarity change reflected the structure of the task, by correlating it with the time elapsed between objects pairs (top right matrix shows median elapsed time between object encounters along the route averaged across participants). For each participant, we compared the correlation between pattern similarity change and the prediction matrix to a surrogate distribution obtained via bootstrapping and used the resulting z-statistic for group-level analysis (see Material and methods).

RESULTS

We examined the effect of learning temporal and spatial positions of objects along a route through a virtual city (Figure 1). Specifically, we presented object images twelve times before and after learning, using the same random order in both scanning runs (Figure 1—figure supplement 1). Taking advantage of the dissociation of temporal and spatial structure through the use of teleporters, we compared the change in neural pattern similarity of entorhinal object representations to the temporal and spatial structure of the task. The change in multi-voxel pattern similarity in alEC between pre- and post-learning scans was negatively correlated with temporal distances between objects pairs along the route (Figure 2A and Figure 2B, $T(25)=-3.75$, $p=0.001$, alpha-level of 0.0125, Bonferroni-corrected for four comparisons). After relative to before learning, objects encountered in temporal proximity were represented more similarly compared to object pairs further separated in time (Figure 2C). Here, we quantified the temporal structure of the task using the median time elapsed between object encounters, but — since object locations were spread regularly along the route — operationalizing the temporal structure in terms of the ordinal distances between the object positions in the sequence yielded comparable results ($T(25)=-3.37$, $p=0.002$; Figure 2—figure supplement 1A).

Pattern similarity change in alEC did not correlate significantly with spatial distances ($T(25)=0.81$, $p=0.420$) and pattern similarity change in posterior-medial EC (pmEC) did not correlate with spatial ($T(25)=0.58$, $p=0.583$) or temporal ($T(25)=1.73$, $p=0.089$) distances. Temporal distances between objects during the first picture viewing task were not related to alEC pattern similarity change (Figure 2—figure supplement 1C; $T(24)=-0.29$ $p=0.776$, one outlier excluded, see Material and methods) and correlations with elapsed time between objects during navigation were significantly more negative ($T(24)=-1.76$ $p=0.045$; one-sided test); strengthening our interpretation that pattern similarity changes reflected relationships experienced in the virtual city.

Can we reconstruct the timeline of events from pattern similarity change in alEC? Here, we used multidimensional scaling to extract coordinates along one dimension from pattern similarity change averaged across participants (Figure 3A-D). The reconstructed temporal coordinates, transformed into the original value range using Procrustes analysis (Figure 3A), mirrored the time points at which objects were encountered during the task (Figure 3B, Pearson correlation between reconstructed and true time points, $r=0.56$, $p=0.023$, bootstrapped 95% confidence interval:

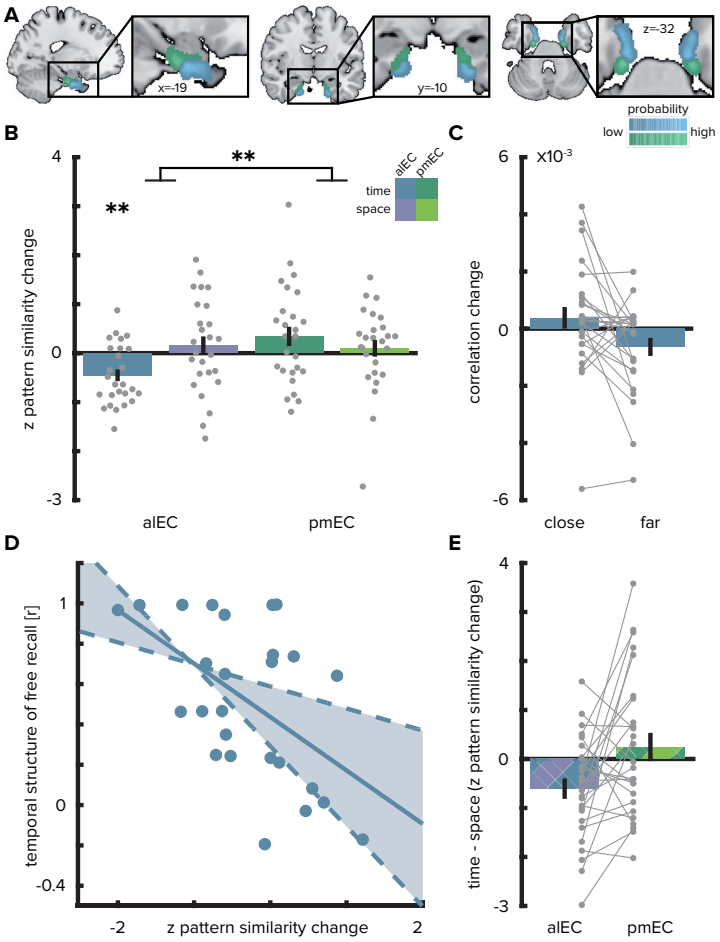


Figure 2 | Temporal mapping in aIEC.

A. Entorhinal cortex subregion masks from (Navarro Schröder et al., 2015) were moved into subject-space and intersected with participant-specific Freesurfer parcellations of entorhinal cortex. Color indicates probability of voxels to belong to the aIEC (blue) or pmEC (green) subregion mask after subject-specific masks were transformed back to MNI template space for visualization. **B.** Pattern similarity change in the aIEC reflected the temporal structure of object encounters along the route as indicated by z-statistics significantly below 0. A permutation-based two-way repeated measures ANOVA further revealed a significant interaction highlighting a difference in temporal and spatial mapping between aIEC and pmEC. **C.** Pattern similarity change in aIEC for objects encountered close together or far apart in time along the route. **D.** Pattern similarity change in aIEC was negatively related to temporal relationships independent of objects encountered in direct succession (Figure 2—figure supplement 1B). The magnitude of this effect correlated significantly with participants' free recall behavior. The temporal organization of freely recalled objects was assessed by calculating the absolute difference in position for all recalled objects and correlating this difference with the time elapsed between encounters of object pairs along the route. Solid line shows least squares line; dashed lines and shaded region highlight bootstrapped confidence intervals. **E.** To illustrate the interaction effect shown in **B**, the difference in the relationship between temporal and spatial distances to pattern similarity change is shown for aIEC and pmEC. Negative values indicate stronger correlations with temporal compared to spatial distances. Bars show mean and S.E.M with lines connecting data points from the same participant in **C** and **E**.

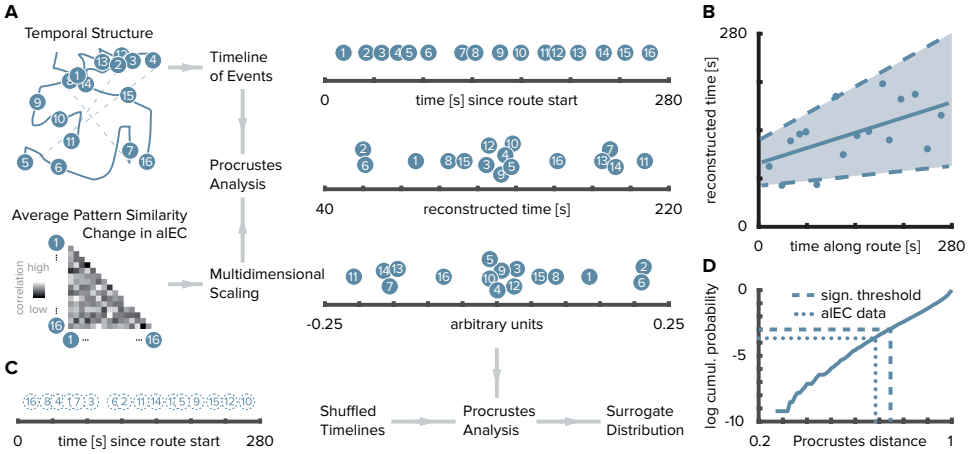


Figure 3 | Reconstructing the timeline of events from entorhinal pattern similarity change.

A. To recover the temporal structure of events we performed multidimensional scaling on the average pattern similarity change matrix in alEC. The resulting coordinates, one for each object along the route, were subjected to Procrustes analysis, which applies translations, rotations and uniform scaling to superimpose the coordinates from multidimensional scaling on the true temporal coordinates along the route (see Material and methods). For visualization, we varied the positions resulting from multidimensional scaling and Procrustes analysis along the y-axis. **B.** The temporal coordinates of this reconstructed timeline were significantly correlated with the true temporal coordinates of object encounters along the route. Circles indicate time points of object encounters; solid line shows least squares line; dashed lines and shaded region highlight bootstrapped confidence intervals. **C.** The goodness of fit of the reconstruction (the Procrustes distance) was compared to a surrogate distribution of Procrustes distances obtained from randomly shuffling the true coordinates against the coordinates obtained from multidimensional scaling and then performing Procrustes analysis for each of 10000 shuffles (left shows one randomly shuffled timeline for illustration). **D.** The Procrustes distance obtained from fitting to the true timeline of events (dotted line) was smaller than the 5th percentile (dashed line) of the surrogate distribution (solid line), which constitutes the significance threshold at an alpha level of 0.05.

0.21, 0.79). Further, we contrasted the fit of the coordinates from multidimensional scaling between the true and randomly shuffled timelines (Figure 3C). Specifically, we compared the deviance of the fit between the reconstructed and the true timeline, the Procrustes distance, to a surrogate distribution of Procrustes distances. This surrogate distribution was obtained by fitting the coordinates from multidimensional scaling to randomly shuffled timelines of events. The Procrustes distance from fitting to the true timeline was smaller than the 5th percentile of the surrogate distribution generated via 10000 random shuffles (Figure 3D, p=0.026). Taken together, these findings indicate that alEC representations change through learning to reflect the temporal structure of the acquired event memories and that we can recover the timeline of events from this representational change.

What is the nature of regional specificity within entorhinal cortex? In a next step, we compared temporal and spatial mapping between the subregions of the entorhinal

cortex. We conducted a permutation-based two-by-two repeated measures ANOVA (see Material and methods) with the factors entorhinal subregion (alEC vs. pmEC) and relationship type (temporal vs. spatial distance between events). Crucially, we observed a significant interaction between EC subregion and distance type ($F(1,25)=7.40$, $p=0.011$, Figure 2B and 2E). Further, the main effect of EC subregion was significant ($F(1,25)=5.18$, $p=0.029$), while the main effect of distance type was not ($F(1,25)=0.84$, $p=0.367$). Based on the significant interaction, we conducted planned post-hoc comparisons, which revealed significant differences (Bonferroni-corrected alpha-level of 0.025) between the mapping of temporal and spatial distances in alEC ($T(25)=-2.91$, $p=0.007$) and a significant difference between temporal mapping in alEC compared to pmEC ($T(25)=-3.52$, $p=0.001$). Spatial and temporal signal-to-noise ratios did not differ between alEC and pmEC (Figure 2—figure supplement 2), ruling out that differences in signal quality might explain the observed effects. Collectively, these findings demonstrate that, within the EC, only representations in the anterior-lateral subregion change to resemble the temporal structure of events and that this mapping was specific to the temporal rather than the spatial dimension.

DISCUSSION

We examined the similarity of multi-voxel patterns to demonstrate that alEC representations reflect the experienced temporal event structure. Despite being cued in random order after learning, these representations related to a holistic temporal map of the task structure. Moreover, entorhinal pattern similarity related to participants' recall behavior and we recovered the timeline of events during learning from these representations.

Our hypothesis for temporal mapping in the alEC was based on a recent finding demonstrating that population activity in the rodent lateral EC carries information from which time can be decoded at different scales ranging from seconds to days (Tsao et al., 2018). Time could be decoded with higher accuracies from the lateral EC than the medial EC and hippocampal subfield CA3. During a structured task in which the animal ran repeated laps on a maze separated into different trials, neural trajectories through population activity space were similar across trials, illustrating that the dynamics of lateral EC neural signals were more stable than during free foraging (Tsao et al., 2018). Consistently, temporal coding was improved for time within a trial during the structured task compared to episodes of free foraging.

These findings support the notion that temporal information in the lateral EC might inherently arise from the encoding of experience (Tsao et al., 2018). In our task, relevant factors contributing to a similar experience of the route on each lap are not only the encounters of objects in a specific order at their respective positions, but also recognizing and passing salient landmarks as well as navigational demands in general. Changes in metabolic states and arousal presumably varied more linearly over time. Slowly drifting activity patterns have been observed also in the human medial temporal lobe (Folkerts et al., 2018) and EC specifically (Lositsky et al., 2016). A representation of time within a known trajectory in the alEC could underlie the encoding of temporal relationships between events in our task, where participants repeatedly navigated along the route to learn the positions of objects. Hence, temporal mapping in the alEC as we report here might help integrate hippocampal spatio-temporal event maps (Deuker et al., 2016).

In a recent fMRI study, in which participants indicated when a still frame was encountered over the course of an episode of a sitcom, the alEC activated more strongly for the third of trials in which participants recalled temporal information most accurately compared to the third of trials in which temporal precision was lowest. These findings implicate the alEC in the retrieval of temporal information, but did not speak to the content of alEC mnemonic representations. Our findings demonstrate that alEC representations reflect the temporal structure of events after learning. One possibility for why alEC multi-voxel patterns resemble a holistic temporal map of the event structure is the reactivation of temporal context information. If alEC neural populations traverse similar population state trajectories on each lap, they would carry information about time within a lap. A given object would be associated with a similar alEC population state on each lap. Associations with temporally drifting signals during the learning task would result in representational changes relative to the baseline scan that, if reactivated in the post-learning picture viewing task, reflect the experienced temporal structure of object encounters. This might explain the observed pattern similarity structure with relatively increased similarity for objects encountered in temporal proximity during learning and decreased similarity for items encountered after longer delays. While this interpretation is in line with data from rodent electrophysiology (Tsao et al., 2018) and the framework proposed by the temporal context model (Howard and Kahana, 2002; Howard et al., 2005), we cannot test the reinstatement of specific activity patterns from the learning phase directly since fMRI data were only collected during the picture viewing tasks in this study. Going beyond the relationship of univariate activation differences to temporal memory precision, we focused on the content of alEC activation patterns and

demonstrate that the alEC represents the temporal structure of events after learning.

Importantly, the highly-controlled design of our study supports the interpretation that alEC representations change through learning to map the temporal event structure. The order of object presentations during the scanning sessions was randomized and thus did not reflect the order in which objects were encountered during the learning task. Since the assignment of objects to positions was randomized across participants and we analyzed pattern similarity *change* from a baseline scan, our findings do not go back to prior associations between the objects, but reflect information learned over the course of the experiment. Further, we presented the object images during the scanning sessions not only in the same random order, but also with the same presentation times and inter-stimulus intervals; thereby ruling out that the effects we observed relate to temporal autocorrelation of the BOLD-signal. Taken together, the high degree of experimental control of our study supports the conclusion that alEC representations change to reflect the temporal structure of acquired memories.

The long time scales of lateral EC temporal codes differ from the observation of time cells in the hippocampus and medial EC, which fire during temporal delays in highly trained tasks (Eichenbaum, 2014; Heys and Dombeck, 2018; Kraus et al., 2015; MacDonald et al., 2011; Mau et al., 2018; Pastalkova et al., 2008). Time cell ensembles change over minutes and days (Mau et al., 2018), but their firing has been investigated in the context of short delays in the range of seconds. This leaves open the questions whether and how time cell sequences encode longer temporal intervals as in our task, where one lap of the route took approximately five minutes.

Our assessment of temporal representations in the antero-lateral and posterior-medial subdivision of the EC was inspired by a recent report of temporal coding during free foraging and repetitive behavior in the rodent EC, which was most pronounced in the lateral EC (Tsao et al., 2018). In humans, local and global functional connectivity patterns suggest a preserved bipartite division of the EC, but along not only its medial-lateral, but also its anterior-posterior axis (Navarro Schröder et al., 2015; Maass et al., 2015). Via these entorhinal subdivisions, cortical inputs from the anterior-temporal and posterior-medial memory systems might converge onto the hippocampus (Ranganath and Ritchey, 2012; Ritchey et al., 2015).

Our findings are in line with the role of the hippocampus in the retrieval of temporal information from memory (Copara et al., 2014; DuBrow and Davachi, 2014; Kyle et al., 2015; Nielson et al., 2015). Hippocampal pattern similarity has been shown

to scale with temporal distances between events (Deuker et al., 2016; Nielson et al., 2015) and evidence for the reinstatement of temporally associated items from memory has been reported in the hippocampus (DuBrow and Davachi, 2014). Already at the stage of encoding, hippocampal and entorhinal activity have been related to later temporal memory (DuBrow and Davachi, 2014, 2016; Ezzyat and Davachi, 2014; Jenkins and Ranganath, 2010, 2016; Lositsky et al., 2016; Tubridy and Davachi, 2011). For example, increased pattern similarity has been reported for items remembered to be close together compared to items remembered to be far apart in time, despite the same time having elapsed between these items (Ezzyat and Davachi, 2014). Similarly, changes in EC pattern similarity during the encoding of a narrative correlated with later duration estimates between events (Lositsky et al., 2016). Complementing these reports, our findings demonstrate that entorhinal activity patterns carry information about the temporal structure of memories at retrieval. Furthermore, the degree to which EC patterns reflected holistic representations of temporal relationships related to recall behavior characterized by the consecutive reproduction of objects encountered in temporal proximity; potentially through mental traversals of the route during memory recall. The central role of the hippocampus and entorhinal cortex in temporal memory (for review see (Davachi and DuBrow, 2015; Howard, 2018; Ranganath, 2018; Wang and Diana, 2017)) dovetails with the involvement of these regions in learning sequences and statistical regularities in general (Barnett et al., 2014; Garvert et al., 2017; Hsieh et al., 2014; Kumaran and Maguire, 2006; Schapiro et al., 2012, 2016; Thavabalasingam et al., 2018)).

In conclusion, our findings demonstrate that activity patterns in alEC, the human homologue region of the rodent lateral EC, carry information about the temporal structure of experienced events. The observed effects might be related to the reactivation of temporal contextual tags, in line with the recent report of temporal information available in rodent lateral EC population activity and models of episodic memory.

MATERIALS AND METHODS

Subjects

26 participants (mean±std. 24.88±2.21 years of age, 42.3% female) were recruited via the university's online recruitment system and participated in the study. As described in the original publication using this dataset (Deuker et al., 2016), this

sample size was based on a power-calculation (alpha-level of 0.001, power of 0.95, estimated effect size of d=1.03 based on a prior study (Milivojevic et al., 2015)) using G*Power (http://www.gpower.hhu.de/). Participants with prior knowledge of the virtual city (see Deuker et al. (2016)) were recruited for the study. All procedures were approved by the local ethics committee (CMO Regio Arnhem Nijmegen) and all participants gave written informed consent prior to commencement of the study.

Design

Overview

The experiment began by a 10 minute session during which participants freely navigated the virtual city (Bellmund et al., 2018b) on a computer screen to re-familiarize themselves with its layout. Afterwards participants were moved into the scanner and completed the first run of the picture viewing task during which they viewed pictures of everyday objects as described below (Figure 1—figure supplement 1). After this baseline scan, participants learned a fixed route through the virtual city along which they encountered the objects at predefined positions (Figure 1 and Figure 1—figure supplement 1). The use of teleporters, which instantaneously moved participants to a different part of the city, enabled us to dissociate temporal and spatial distances between object positions (Figure 1—figure supplement 2). Subsequent to the spatio-temporal learning task, participants again underwent fMRI and completed the second run of the picture viewing task. Lastly, participants' memory was probed outside of the MRI scanner. Specifically, participants freely recalled the objects they encountered, estimated spatial and temporal distances between them on a subjective scale, and indicated their knowledge of the positions the objects in the virtual city on a top-down map (Deuker et al., 2016).

Spatio-temporal learning task

Participants learned the positions of everyday objects along a trajectory through the virtual city Donderstown (Bellmund et al., 2018b). This urban environment, surrounded by a range of mountains, consists of a complex street network, parks and buildings. Participants with prior knowledge of the virtual city (see Deuker et al. (2016)) were recruited for the study. After the baseline scan, participants navigated the fixed route through the city along which they encountered 16 wooden chests at specified positions (Figure 1A). During the initial 6 laps the route was marked by traffic cones. In later laps, participants had to rely on their memory to navigate the route, but guidance in the form of traffic cones was available upon button press for laps 7-11. Participants completed 14 laps of the route in total (mean ± standard

deviation of duration 71.63±13.75 minutes), which were separated by a black screen displayed for 15s before commencement of the next lap from the start position.

Participants were instructed to open the chests they encountered along the route by walking into them. They were then shown the object contained in that chest for 2 seconds on a black screen. A given chest always contained the same object for a participant, with the assignment of objects to chests randomized across participants. Therefore, each object was associated with a spatial position defined by its location in the virtual city and a temporal position described by its occurrence along the progression of the route. Importantly, we dissociated temporal relationships between object pairs (measured by time elapsed between their encounter) from the Euclidean distance between their positions in the city through the use of teleporters. Specifically, at three locations along the route participants encountered teleporters, which immediately transported them to a different position in the city where the route continued (Figure 1A). This manipulation allows the otherwise impossible encounter of objects after only a short temporal delay, but with a large Euclidean distance between them in the virtual city (Deuker et al., 2016). Indeed, temporal and spatial distances across all comparisons of object pairs were uncorrelated (Pearson r=-0.068; bootstrapped 95% confidence interval: -0.24, 0.12; p=0.462; Figure 1—figure supplement 2).

Picture viewing tasks

Before and after the spatio-temporal learning task, participants completed the picture viewing tasks while undergoing fMRI (Deuker et al., 2016). During these picture viewing tasks, the 16 objects from the learning task as well as an additional target object were presented. Participants were instructed to attend to the objects and to respond via button press when the target object was presented. Every object was shown 12 times in 12 blocks, with every object being shown once in every block. In each block, the order of objects was randomized. Blocks were separated by a 30 second break without object presentation. Objects were presented for 2.5 seconds on a black background in each trial and trials were separated by two or three TRs. These intertrial intervals occurred equally often and were randomly assigned to the object presentations. The presentation of object images was locked to the onset of the new fMRI volume. For each participant, we generated a trial order adhering to the above constraints and used the identical trial order for the picture viewing tasks before and after learning the spatio-temporal arrangement of objects along the route. Using the exact same temporal structure of object presentations in both runs rules out potential effects of temporal autocorrelation of the BOLD signal on

the results, since such a spurious influence on the representational structure would be present in both tasks similarly and therefore cannot drive the pattern similarity *change* we focused our analysis on (Deuker et al., 2016).

MRI Acquisition

All MRI data were collected using a 3T Siemens Skyra scanner (Siemens, Erlangen, Germany). Functional images during the picture viewing tasks were acquired with a 2D EPI sequence (voxel size 1.5mm isotropic, TR=2270ms, TE=24ms, 40 slices, distance factor 13%, flip angle 85°, field of view (FOV) 210×210×68mm). The FOV was oriented to fully cover the medial temporal lobes and if possible calcarine sulcus (Deuker et al., 2016). To improve the registration of the functional images with partial coverage of the brain, 10 volumes of the same functional sequence with an increased number of slices (120 slices, TR=6804.1ms) were acquired (see fMRI preprocessing). Additionally, gradient field maps were acquired (for 21 participants) with a gradient echo sequence (TR=1020 ms; TE1=10ms; TE2=12.46ms; flip angle=90˚; volume resolution=3.5×3.5×2mm; FOV = 224×224 mm). Further, a structural image was acquired for each participant (voxel size = 0.8×0.8×0.8mm, TR=2300ms; TE=315ms; flip angle=8˚; in-plane resolution=256×256 mm; 224 slices).

Behavioral Data

Results from in-depth analysis of the behavioral data obtained during the spatio-temporal learning task as well as the memory tests conducted after fMRI scanning are reported in detail in Deuker et al. (2016). Here, we used data from the spatio-temporal learning task as predictions for multi-voxel pattern similarity (see below). Specifically, we defined the temporal structure of pairwise relationships between objects pairs as the median time elapsed between object encounters across the 14 laps of the route. These times differed between participants due to differences in navigation speed (Deuker et al., 2016). Figure 1b shows the temporal distance matrix averaged across participants for illustration. In our task, chests containing objects were spread evenly along the route and hence ordinal distances between objects provide a closely related measure of temporal structure (mean±standard deviation Pearson r=0.993±0.0014, Figure 2—figure supplement 1A). The spatial distances between object positions were defined as the Euclidean distances between the locations of the respective chests in the virtual city. For details of the analysis quantifying the relationship between entorhinal pattern similarity change and recall behavior see the corresponding section below.

fMRI preprocessing

Preprocessing of fMRI data was carried out using FEAT (FMRI Expert Analysis Tool, version 6.00), part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl, version 5.0.8), as described in (Deuker et al., 2016). Functional images were submitted to motion correction and high-pass filtering (cutoff 100s). Images were not smoothed. When available, distortion correction using the fieldmaps was applied. Using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002), the functional images acquired during the picture viewing tasks were registered to the preprocessed whole-brain mean functional images, which were in turn registered to the to the participant's structural scan. The linear registration from this high-resolution structural to standard MNI space (1mm resolution) was then further refined using FNIRT nonlinear registration (Anderson et al., 2010). Representational similarity analysis of the functional images acquired during the picture viewing tasks was carried out in regions of interests co-registered to the space of the whole-brain functional images.

ROI definition

Based on functional connectivity patterns, the anterior-lateral and posterior-medial portions of human EC were identified as human homologue regions of the rodent lateral and medial EC in two independent studies (Navarro Schröder et al., 2015; Maass et al., 2015). Here, we focused on temporal coding in the alEC, building upon a recent report of temporal signals in rodent lateral EC during navigation (Tsao et al., 2018). Therefore, we used masks from (Navarro Schröder et al., 2015) to perform ROI-based representational similarity analysis on our data. For each ROI, the mask was co-registered from standard MNI space (1mm) to each participant's functional space (number of voxels: alEC 126.7±46.3; pmEC 69.0±32.9). To improve anatomical precision for the EC masks, the subregion masks from (Navarro Schröder et al., 2015) were each intersected with participant-specific EC masks obtained from their structural scan using the automated segmentation implemented in Freesurfer (version 5.3).

Representational Similarity Analysis

As described in Deuker et al. (2016), we implemented representational similarity analysis (RSA, Kriegeskorte et al. (2008b, 2008a)) for the two picture viewing tasks individually and then analyzed changes in pattern similarity between the two picture viewing tasks, which were separated by the spatio-temporal learning phase. After preprocessing, analyses were conducted in Matlab (version 2017b, MathWorks). In a general linear model, we used the motion parameters obtained during preprocessing as predictors for the time series of each voxel in the respective

ROI. Only the residuals of this GLM, i.e. the part of the data that could not be explained by head motion, were used for further analysis. Stimulus presentations during the picture viewing tasks were locked to the onset of fMRI volumes and the third volume after the onset of picture presentations, corresponding to the time 4.54 to 6.81 seconds after stimulus onset, was extracted for RSA.

For each ROI, we calculated Pearson correlation coefficients between all object presentations except for comparisons within the same of the 12 blocks of each picture viewing task. For each pairwise comparison, we averaged the resulting correlation coefficients across comparisons, yielding a 16×16 matrix reflecting the average representational similarity of objects for each picture viewing task (Deuker et al., 2016). These matrices were Fisher z-transformed. Since the picture viewing task was conducted before and after spatio-temporal learning, the two cross-correlation matrices reflected representational similarity with and without knowledge of the spatial and temporal relationships between objects, respectively. Thus, the difference between the two matrices corresponds to the change in pattern similarity due to learning. Specifically, we subtracted the pattern similarity matrix obtained prior to learning from the pattern similarity matrix obtained after learning, resulting in a matrix of pattern similarity change for each ROI from each participant. This change in similarity of object representations was then compared to different predictions of how this effect of learning might be explained (Figure 1B).

To test the hypothesis that multi-voxel pattern similarity change reflects the temporal structure of the object encounters along the route, we correlated pattern similarity change with the temporal relationships between object pairs; defined by the participant-specific median time elapsed between object encounters while navigating the route. Likewise, we compared pattern similarity change to the Euclidean distances between object positions in the virtual city as well as the temporal relations subjectively remembered by each participant. We calculated Spearman correlation coefficients to quantify the fit between pattern similarity change and each prediction. We expected negative correlations as relative increases in pattern similarity are expected for objects separated by only a small distance compared to comparisons of objects separated by large distances (Deuker et al., 2016). We compared these correlation coefficients to a surrogate distribution obtained from shuffling pattern similarity change against the respective prediction. For each of 10000 shuffles, the Spearman correlation coefficient between the two variables was calculated, yielding a surrogate distribution of correlation coefficients (Figure 1B). We quantified the size of the original correlation coefficient in comparison to the surrogate distribution.

Specifically, we assessed the proportion of larger or equal correlation coefficients in the surrogate distribution and converted the resulting p-value into a z-statistic using the inverse of the normal cumulative distribution function (Deuker et al., 2016; Stelzer et al., 2013; Schlichting et al., 2015). Thus, for each participant, we obtained a z-statistic reflecting the fit of the prediction to pattern similarity change in that ROI. For visualization (Figure 2c), we averaged correlation coefficients quantifying pattern similarity change in alEC separately for comparisons of objects encountered close together or far apart in time based on the median elapsed time between object pairs.

The z-statistics were tested on the group level using permutation-based procedures (10000 permutations) implemented in the Resampling Statistical Toolkit for Matlab (https://mathworks.com/matlabcentral/fileexchange/27960-resampling-statistical-toolkit). To test whether pattern similarity change in alEC reflected the temporal structure of object encounters, we tested the respective z-statistic against 0 using a permutation-based t-test and compared the resulting p-value against an alpha of 0.0125 (Bonferroni-corrected for 4 comparisons, Figure 2). Respecting within-subject dependencies, differences between the fit of temporal and spatial relationships between objects and pattern similarity change in the EC subregions were assessed using a permutation-based two-way repeated measures ANOVA with the factors EC subregion (alEC vs. pmEC) and relationship type (elapsed time vs. Euclidean distance). Planned post-hoc comparisons then included permutation-based t-tests of temporal against spatial mapping in alEC and temporal mapping between alEC and pmEC (Bonferroni-corrected alpha-level of 0.025).

Accounting for adjacency effects

To rule out that only increased pattern similarity for object pairs encountered at adjacent temporal positions along the route drove the effect we excluded these comparisons from the analysis when testing whether pattern similarity change in alEC reflected temporal relationships. We tested the resulting z-values, reflective of holistic temporal maps independent of direct adjacency, against 0 (Figure 2—figure supplement 1B, one-sided permutation-based t-test). The z-values of this analysis were used for the correlation with recall behavior described below and shown in Figure 2D.

Relationship between pattern similarity change and recall behavior

We assessed participants' tendency to reproduce objects encountered closely in time along the route at nearby positions during free recall. In this task, conducted

after the post-learning picture viewing task, participants had two minutes to name as many of the objects encountered in the virtual city as possible and to speak the names in the order in which they came to mind into a microphone (Deuker et al., 2016). For each pair of recalled objects we calculated the absolute positional difference in reproduction order and correlated these recall distances with elapsed time between object encounters of these pairs. This resulted in high Pearson correlation coefficients for participants with the tendency to recall objects at distant temporal positions along the route far apart and to retrieve objects encountered closely together in time along the route at nearby positions during memory retrieval. Such a temporally structured recall order would result for example from mentally traversing the route during the free recall task. The temporal structure of participants' free recall was significantly correlated with the strength of the relationship between elapsed time and pattern similarity change after excluding comparisons of objects encountered at directly adjacent temporal positions (Figure 2D).

Temporal intervals during the baseline scan

We interpret pattern similarity change between the picture viewing tasks as being induced by the learning task. To rule out effects of temporal intervals between objects experienced outside of the virtual city we correlated pattern similarity change in the alEC with temporal relationships during the pre-learning baseline scan. Specifically, we calculated the average temporal distance during the first picture viewing task for each pair of objects. Analogous to the time elapsed during the task, we correlated these temporal distances with pattern similarity change in the alEC. One participant was excluded from this analysis due to a z-value more than 1.5 times the interquartile range below the lower quartile. We tested whether pattern similarity change differed from zero and whether correlations with elapsed time during the task were more negative than correlations with temporal distance during the first picture viewing task (one-sided test) using permutation-based t-tests (Figure 2—figure supplement 1C).

Timeline reconstruction

To reconstruct the timeline of events from alEC pattern similarity change we combined multidimensional scaling with Procrustes analysis (Figure 2A). We first rescaled the pattern similarity matrix to a range from 0 to 1 and then converted it to a distance matrix (distance = 1-similarity). We averaged the distance matrices across participants and subjected the resulting matrix to classical multidimensional scaling. Since we were aiming to recover the timeline of events, we extracted coordinates underlying the averaged pattern distance matrix along one dimension. In a next

step, we fitted the resulting coordinates to the times of object encounters along the route, which were also averaged across participants, using Procrustes analysis. This analysis finds the linear transformation, allowing scaling and reflections, that minimizes the sum of squared errors between the two sets of temporal coordinates. To assess whether the reconstruction of the temporal relationships between memories was above chance, we correlated the reconstructed temporal coordinates with the true temporal coordinates using Pearson correlation (Figure 2B). 95% confidence intervals were bootstrapped using the Robust Correlation Toolbox (Pernet et al., 2013). Additionally, we compared the goodness of fit of the Procrustes transform—the Procrustes distance, which measures the deviance between true and reconstructed coordinates—to a surrogate distribution. Specifically, we randomly shuffled the true temporal coordinates and then mapped the coordinates from multidimensional scaling onto these shuffled timelines. We computed the Procrustes distance for each of 10000 iterations. We quantified the proportion of random fits in the surrogate distribution better than the fit to the true timeline (i.e. smaller Procrustes distances) and expressed it as a p-value to demonstrate that our reconstruction exceeds chance level (Figure 2C-D).

Signal-to-noise ratio

We quantified the temporal and spatial signal-to-noise ratio for each ROI. Temporal signal-to-noise was calculated for each voxel as the temporal mean divided by the temporal standard deviation for both runs of the picture viewing task separately. Values were averaged across the two runs and across voxels in the ROIs. Spatial signal-to-noise ratio was calculated for each volume as the mean signal divided by the standard deviation across voxels in the ROI. The resulting values were averaged across volumes of the time series and averaged across the two runs. Signal-to-noise ratios were compared between ROIs using permutation-based t-tests.

FIGURE SUPPLEMENTS

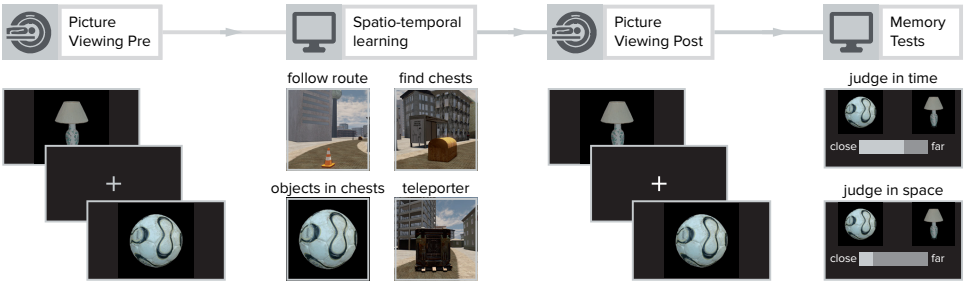


Figure 1 - figure supplement 1 | Overview of experimental design.

Participants viewed object images in random order while undergoing fMRI before and after learning the temporal and spatial relationships between these objects The order and timing of picture presentations was held identical in both sessions to assess changes in the similarity of object representations as measured by the difference in similarity of multi-voxel activity patterns (see Material and methods). In between the two picture viewing tasks, participants acquired knowledge about the spatial and temporal positions of objects along a route through the virtual city. Initially, the route was marked by traffic cones, but in later laps participants navigated the route without guidance. Participants encountered chests along the route and were instructed to open the chests by walking into them. Each chest contained a different object, which was displayed on a black screen upon opening the chest. Crucially, the route featured three teleporters that instantly teleported participants to a different part of the city where the route continued (Figure 1). This manipulation enabled us to dissociate the temporal and spatial distances between pairwise object comparisons (Figure1—figure supplement 2). After the second picture viewing task, participants' memory for temporal and spatial relationships between object pairs was assessed. Here, participants adjusted a slider to indicate whether they remembered object pairs to be close together or far apart. Temporal and spatial relations were judged in separate trials. The results of these memory tests are reported in detail in Deuker et al. (2016).

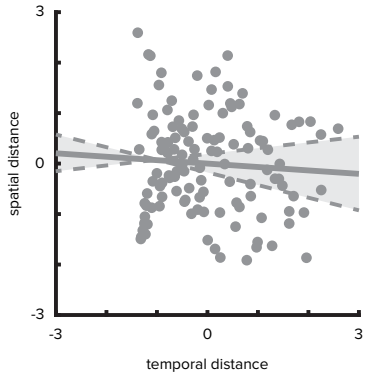


Figure 1 - figure supplement 2 | Temporal and spatial distances are uncorrelated.

Pairwise temporal and spatial distances between objects are uncorrelated (Pearson r=-0.068; bootstrapped 95% confidence interval: -0.24, 0.12; p=0.462). Median times elapsed between object encounters were z-scored and then averaged across participants. Spatial distances were defined as z-scored Euclidean distances between object positions. When correlating individual median times elapsed with spatial distances, the correlation between the dimensions was not significant in any of the participants (mean ± standard deviation of Pearson correlation coefficients r= -0.068±0.006, all p≥0.378).

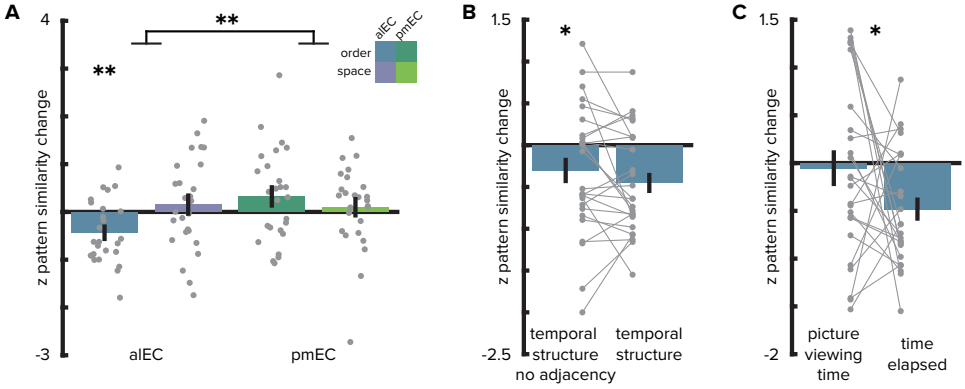


Figure 2 - figure supplement 1 | Control analyses.

A. Repeating the two-way repeated measures ANOVA using ordinal distances as the measure of temporal structure yielded results comparable to the analyses presented in Figure 2. We observed a significant interaction highlighting a difference in temporal and spatial mapping between aIEC and pmEC ($F(1,25)=7.11$, $p=0.012$; post hoc tests comparing mapping of temporal order vs. spatial distance in the aIEC ($T(25)=-2.81$, $p=0.008$) and comparing mapping of temporal order between aIEC and pmEC ($T(25)=-3.53$, $p=0.002$) significant at Bonferroni-corrected alpha-level of 0.025; main effect of EC subregion ($F(1,25)=4.97$, $p=0.033$); main effect of distance type ($F(1,25)=0.84$, $p=0.365$)). **B.** To rule out that the effect was driven by objects at temporally adjacent positions along the route we excluded these comparisons from the analysis. The effect of temporal mapping in the aIEC remained significant ($T(25)=-2.00$, $p=0.029$, one-sided test) and the result of this analysis did not differ from the original results obtained from the analysis including all comparisons ($T(25)=1.40$, $p=0.190$). **C.** Pattern similarity change in aIEC was not correlated with temporal distances from the first picture viewing task. Correlations with pattern similarity change were more negative for elapsed time during the learning task than for presentation times during the first picture viewing task (see Material and methods).

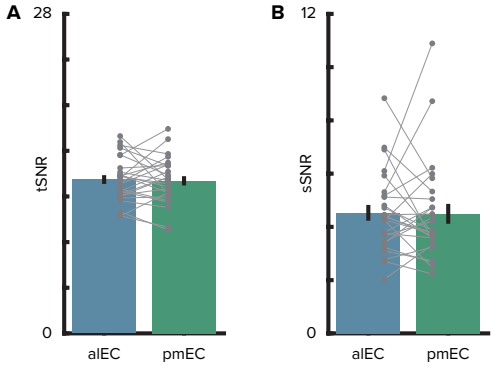
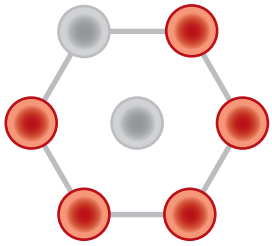


Figure 2 - figure supplement 2 | Signal-to-noise ratio in the entorhinal cortex.

A. The temporal signal-to-noise ratio did not differ between the entorhinal subregions. **B.** Similarly, the spatial signal-to-noise ratio did not differ between entorhinal subregions. Bars show mean and S.E.M with lines connecting data points from the same participant.

CHAPTER 5

Navigating Cognition: Spatial Codes for Human Thinking



This chapter is published as:

J. L. S. Bellmund, P. Gärdenfors, E. I. Moser, C. F. Doeller (2018). Navigating cognition: Spatial codes for human thinking. *Science, 362*, eaat6766. https://doi.org/10.1126/science.aat6766

ABSTRACT

The hippocampal formation has long been suggested to underlie both memory formation and spatial navigation. Here, we discuss how neural mechanisms identified in spatial navigation research operate across information domains to support a wide spectrum of cognitive functions. In our framework, place and grid cell population codes provide a representational format to map variable dimensions of cognitive spaces. This highly dynamic mapping system enables rapid reorganization of codes through remapping between orthogonal representations across behavioral contexts, yielding a multitude of stable cognitive spaces at different resolutions and hierarchical levels. Action sequences result in trajectories through cognitive space, which can be simulated via sequential coding in the hippocampus. Thereby, the spatial representational format of the hippocampal formation has the capacity to support flexible cognition and behavior.

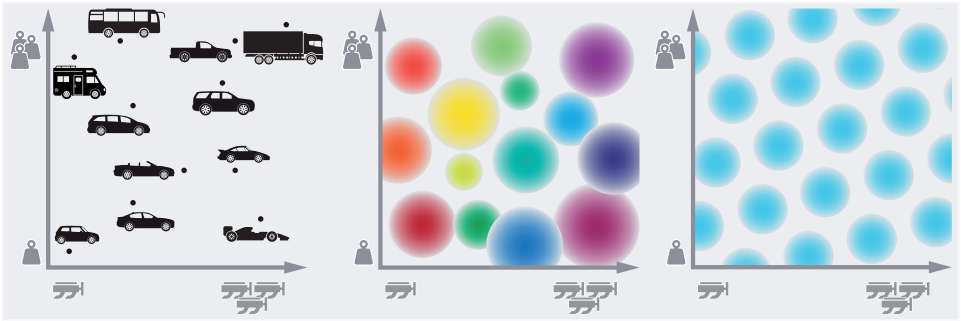
REVIEW SUMMARY

Background

Ever since Edvard Tolman's proposal that comprehensive cognitive maps underlie spatial navigation and, more generally, psychological functions, the question how past experience guides behavior has been contentious. The discovery of place cells in rodents, signaling the animal's position in space, suggested such cognitive maps to reside in the hippocampus, a core brain region for human memory. Building upon the description of place cells, several other functionally defined cell types were discovered in the hippocampal-entorhinal region. Among them are grid cells in the entorhinal cortex, whose characteristic periodic, six-fold symmetric firing patterns are thought to provide a spatial metric. These findings were complemented by insights into key coding principles of the hippocampal-entorhinal region: Spatial representations vary in scale along the hippocampal long axis, place cells remap to map different environments, and sequential hippocampal activity represents nonlocal trajectories through space. In humans, the existence of spatially tuned cells has been demonstrated in presurgical patients and fMRI provides proxy-measures for the non-invasive investigation of these processing mechanisms in human cognition. Intriguingly, recent advances indicate that place and grid cells can encode positions along dimensions of experience beyond Euclidean space for navigation, suggesting a more general role of hippocampal-entorhinal processing mechanisms in cognition.

Advances

We combine hippocampal-entorhinal processing mechanisms identified in spatial navigation research with ideas from cognitive science describing a spatial representational format for cognition. Cognitive spaces are spanned by dimensions satisfying geometric constraints such as betweenness and equidistance, enabling the representation of properties and concepts as convex regions of cognitive space. We propose that the continuous population code of place and grid cells in the hippocampal-entorhinal region maps the dimensions of cognitive spaces. In these, each stimulus is located according to its feature values along the relevant dimensions, resulting in nearby positions for similar stimuli and larger distances between dissimilar stimuli. The low-dimensional, rigid firing properties of the entorhinal grid system make it a candidate to provide a metric or distance code for cognitive spaces, whereas hippocampal place cells flexibly represent positions in a given space. This mapping of cognitive spaces is complemented by the additional coding principles outlined above: Along the hippocampal long-axis cognitive spaces are mapped with varying spatial scale, supporting memory and knowledge representations at



Summary Figure | Place and grid cells map cognitive spaces.

Cognitive spaces are defined by dimensions satisfying geometric constraints. The example space (left) is spanned by the dimensions of engine power and car weight. Black dots show different vehicles whose positions reflect their feature combinations. Place cells (center, colored circles represent firing fields of different cells) and grid cells (right, circles illustrate firing pattern of one cell) provide a continuous code for cognitive spaces.

different levels of granularity. Via hippocampal remapping, spaces spanned by different dimensions can be flexibly mapped and established maps can be reinstated via attractor dynamics. The geometric definition of cognitive spaces allows flexible generalization and inference and sequential hippocampal activity can simulate trajectories through cognitive spaces for adaptive decision-making and behavior.

Outlook

Cognitive spaces provide a domain-general format for processing in the hippocampal-entorhinal region, in line with its involvement beyond navigation and memory. Spatial navigation serves as a model system to identify key coding principles governing cognitive spaces. An important question concerns the extent to which firing properties of spatially tuned cells are preserved in cognitive spaces. Technological advances such as calcium imaging will clarify coding principles on the population level and facilitate the translation to human cognitive neuroscience. Spatial navigation is mostly investigated in two dimensions and naturally limited to three dimensions; however, the processing of complex, multi-dimensional concepts is vital to high-level human cognition and the representation of such high-dimensional spaces is an intriguing question for future research. Further, the role of brain networks acting in concert with the hippocampus in navigation specifically and cognitive function in general will provide insight into whether and how cognitive spaces are supported beyond the hippocampal-entorhinal region. Finally, the precise way in which cognitive spaces and trajectories through them are read out for behavior remains to be elucidated.

INTRODUCTION

How past experience guides behavior has been a highly contested topic for decades. In 1948, Tolman (1948) described evidence for learning beyond stimulus-response driven behavior in rats, foreshadowing the cognitive revolution to follow. He proposed that rats learn comprehensive maps of their environments, which can guide flexible goal-directed behavior such as finding shortcuts. Tolman coined the term *cognitive map* and speculated how these maps might underlie psychological functions. Support for his controversial speculations was provided by the first reports of hippocampal place cells more than 20 years later (O'Keefe and Dostrovsky, 1971). Place cells are generally only active when an animal occupies the cell's preferred location. These cells were proposed by O'Keefe and Nadel (1978) as a potential neural substrate of Tolman's cognitive map. Incorporating the discovery of place cells and findings from lesion studies, the hippocampus was suggested to provide a map-like reference system, signaling both the position of the agent and features of the environment for navigational and mnemonic processing of events (O'Keefe and Nadel, 1978). The sensitivity of hippocampal coding to items or discrete events in the environment (Gothard et al., 1996a; Wood et al., 1999; Young et al., 1994), differential coding of overlapping path segments based on task state (Wood et al., 2000) and the widespread deficits in patients with hippocampal lesions (Cohen and Eichenbaum, 1993) have been captured in the development of relational memory theory (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 1988, 2014). This account builds on the capabilities of the hippocampus for associative binding to link event representations into relational networks and has offered a counterpole to an exclusively spatial processing view of the hippocampus. Episodic memories can be formed by linking successive event representations and episodic trajectories into ensemble patterns stored in hippocampal networks for subsequent retrieval (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 2014; Marr, 1971; Hasselmo, 2011; McNaughton and Morris, 1987; Schiller et al., 2015). Recent advances demonstrate a hippocampal involvement in flexible cognition beyond the domains of navigation and memory (Hannula and Duff, 2017).

SPACE CODES AS A REPRESENTATIONAL FORMAT FOR COGNITION

We here describe domain-general core coding principles from spatial navigation research that have the potential to support a wider span of cognitive functions. Specifically, we propose that the hippocampal-entorhinal system represents

experience in cognitive spaces (cf. conceptual spaces in (Balkenius and Gärdenfors, 2016; Gärdenfors, 2000)). A cognitive space is thought to be spanned by a set of quality dimensions, which can be closely related to sensory inputs but also comprise abstract features (Gärdenfors, 2000). A given stimulus can be located in a cognitive space based on its set of feature values along a set of quality dimensions. Each dimension is equipped with an underlying metric and follows geometric constraints satisfying the mathematical notions of betweenness and equidistance (Gärdenfors, 2000). Consider the simple example that when planning to buy a new car, you might describe cars along two dimensions, their engine power and their weight (Fig. 1). Depending on the two features, vehicles can be located in this two-dimensional space, for instance sports car might be rather lightweight, but have a strong engine. In this example, we might treat the dimensions as separable. However, stimuli can also vary along integral dimensions, on which a given stimulus cannot be described independently. For example, a car's color can only be fully described when assigning values on all three integral dimensions hue, saturation and brightness, which constitute the color domain (Gärdenfors, 2000). Different cognitive spaces might have different underlying metrics, for example the Euclidean or the city-block metric, which allow assigning distances along a dimension reflecting the degree of similarity between locations in cognitive space with similar stimuli located close together (Gärdenfors, 2000; Gärdenfors and Zenker, 2015).

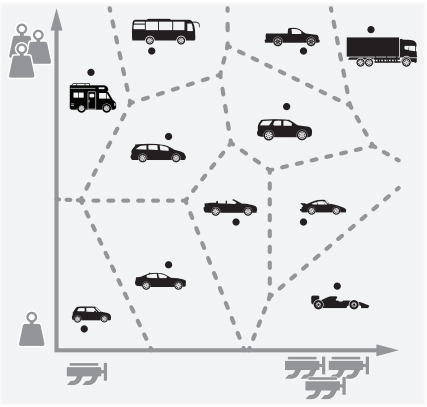


Figure 1 | Two-dimensional cognitive space.

Schematic of a cognitive space spanned by the dimensions of car weight and engine power. Feature values along the two axes define positions of different cars resulting in stimuli with similar properties being located nearby. Concepts (car icons) are defined as convex regions of the cognitive space and are indicated by dashed lines obtained from a Voronoi tessellation of the space. Under the assumption of a Euclidean metric this discretizes the space into convex regions by assigning each point in space to the region around the closest prototypical exemplar (black dots) with distances based on dissimilarity along the feature dimensions.

Building on the geometric characterization of quality dimensions, a property is defined as a convex region in some domain with convexity meaning that for all points x and y in the region the points between x and y also fall into the region (Gärdenfors, 2000). In this framework, a property constitutes the simplest form of a concept, based on only one domain, for example describing a car as 'heavy'. Defining properties as convex regions enables generalization in that a property of two stimuli x and y can be inferred to be shared by any stimulus z falling in between x and y. Therefore, the geometric constraints on cognitive spaces allow inference about never-experienced stimuli. Thereby cognitive spaces afford great cognitive flexibility; going beyond associative and transitive inferences also permitted by relational networks (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 1988, 2014). More complex concepts comprise multiple domains and information about their interrelations. Thus, a concept is defined as a set of convex regions in a number of domains with domains weighted based on salience and additional information about how the regions in different domains are correlated (Gärdenfors, 2000). In our example, racing cars would occupy a region characterized by high power and low weight (Fig. 1). Following the spatial definition of concepts, different exemplars correspond to more or less central positions, with prototypical members (Rosch, 1975) located centrally in the conceptual region. Using a Voronoi tessellation of a continuous space with a Euclidean metric, where all locations in a given space are discretized as belonging to the nearest prototype, convex regions emerge and allow classification of individual stimuli (Gärdenfors, 2000).

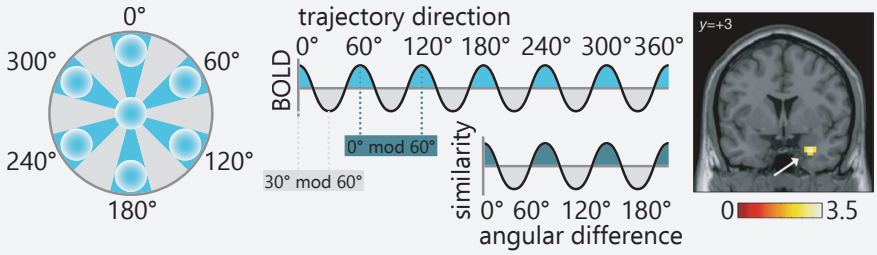
We propose that *processing mechanisms in the hippocampal-entorhinal system are well-suited to support cognitive spaces as a domain-general format for flexible, high-level cognition in humans*. Studies of spatial navigation describe how space is mapped by a continuous, multi-scale code of function-specific cells in the hippocampal-entorhinal system. This system enables flexible mapping of different environments and simulations of trajectories through space via temporally compressed sequences. These processing mechanisms might have developed from originally mapping navigable space to also representing cognitive spaces; consistent with an evolutionary preserved circuitry of the hippocampal-entorhinal region across mammals (Clark and Squire, 2013), the translation of insights from spatial navigation research in rodents to human navigation and beyond (Buzsáki and Moser, 2013; Doeller et al., 2010; O'Keefe and Nadel, 1978) (see Box), and the notion of a core system of geometry (Spelke and Lee, 2012).

FROM SPATIAL NAVIGATION TO COGNITIVE SPACES

In the mammalian brain, positional information is conveyed by the spatially constrained firing of place (O'Keefe and Dostrovsky, 1971) and grid (Hafting et al., 2005) cells during spatial navigation. Place cells in the hippocampus are preferentially active when the animal occupies a certain position within the environment, the cell's receptive field or place field. The firing fields of the population of place cells are thought to cover the entire environment, thereby providing a map-like representation of the animal's surroundings (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978). While the firing of a place cell is usually restricted to one place in a small environment, grid cells in the medial entorhinal cortex (EC), one synapse from the hippocampus, exhibit multiple firing fields located at the vertices of equilateral triangles tiling the entire environment (Hafting et al., 2005). These regular, six-fold symmetric firing patterns are assumed to support spatial navigation by providing a coordinate system of the environment (Hafting et al., 2005; Moser et al., 2017). The entorhinal grid system is assigned a key role in path integration and vector navigation (Bush et al., 2015; Herz et al., 2017; McNaughton et al., 2006; Moser et al., 2017).

Intracranial recordings in patients navigating virtual reality environments established the existence of place (Ekstrom et al., 2003) and grid (Jacobs et al., 2013) cells in humans. The six-fold symmetry of grid-cell firing has been translated to non-invasive functional magnetic resonance imaging (fMRI) in healthy volunteers, where grid-like hexadirectional signals have been observed in EC during navigation (Doeller et al., 2010; Kunz et al., 2015; Constantinescu et al., 2016; Bellmund et al., 2016; Horner et al., 2016; Nau et al., 2018a; Julian et al., 2018; Stangl et al., 2018) (see Box). The brain's spatial navigation system (Moser et al., 2017; Epstein et al., 2017) further includes head direction cells conveying information about the animal's head direction (Cullen and Taube, 2017), goal and goal direction cells signaling egocentric directions to navigational goals (Hok et al., 2005; Sarel et al., 2017) as well as speed cells sensitive to running speed (Kropff et al., 2015) and border (Savelli et al., 2008; Solstad et al., 2008) or boundary vector cells (Lever et al., 2009) responding to borders in the environment.

Box: Hexadirectional signals in fMRI



Hexadirectional signals in fMRI.

The number of grid cell firing fields crossed depends on the direction of a trajectory through cognitive space. More fields are crossed during trajectories aligned with the grid (blue), translating to stronger entorhinal BOLD activity (middle top, see text for details). This effect was first observed during virtual navigation in the entorhinal cortex (right, adapted from Doeller et al., Nature, (2010)). The similarity structure of entorhinal multi-voxel patterns exhibits a 60°-modulation when comparing trajectories as a function of their angular difference as the grid is sampled at the same phase every 60° (middle bottom, red), irrespective of its alignment.

Grid cells are defined by their six-fold symmetric firing patterns tiling environments in a highly regular fashion (Hafting et al., 2005). Hexadirectional signals serve as a proxy measure for grid-like activity in BOLD-fMRI during trajectories through cognitive spaces to investigate the role of the grid system in higher-level cognition in the healthy human brain (Bellmund et al., 2016; Constantinescu et al., 2016; Doeller et al., 2010; Horner et al., 2016; Julian et al., 2018; Kunz et al., 2015; Nau et al., 2018a; Stangl et al., 2018). Three analysis approaches relying on a directional bias of activity have been employed. In the orientation-estimation approach (Doeller et al., 2010) the data are partitioned to estimate the orientation of the hexadirectional signal in one part of the data. The prediction of increased levels of BOLD-activity for trajectories aligned versus misaligned with the estimated orientation is then tested on an independent data partition (see figure). In the second analysis approach, which is based on fMRI repetition suppression, the hexadirectional grid code is reflected in correlations of BOLD-activity during trajectories in a given direction with the time since the last trajectory at an angular offset of 60° (Doeller et al., 2010). The third approach analyzes the similarity of multi-voxel patterns as a function of angular differences between trajectory directions to test the assumption that activity of the grid system should be reflected in a 60°-modulation of entorhinal activity patterns (Bellmund et

al., 2016). Hexadirectional signals can serve as a showcase example for how insights from rodent electrophysiology might be translated to human navigation in fMRI and explored in human cognition and behavior more broadly (Bellmund et al., 2016; Constantinescu et al., 2016; Doeller et al., 2010; Horner et al., 2016; Julian et al., 2018; Kunz et al., 2015; Nau et al., 2018a; Stangl et al., 2018).

While the BOLD-response measured with fMRI does not reflect single cell activity, it provides a hemodynamic proxy of population activity (Logothetis, 2008). How can the 60°-symmetry of grid cell firing patterns be translated to a bias in population activity picked up by fMRI during trajectories through cognitive space? While a single grid cell intuitively exhibits hexadirectional activity, a hexadirectional bias on the population level might result if orientations (Krupic et al., 2015; Stensola et al., 2015) and spatial offsets to boundaries (Heys et al., 2014) of the firing patterns are clustered across cells; which could also potentially be reflected in directional biases in the local field potential in EC; see also recent findings in MEG and intracranial electroencephalography (Staudigl et al., 2018; Chen et al., 2018; Maidenbaum et al., 2018). Additionally, conjunctive grid cells (Moser et al., 2017) also modulated by heading direction might contribute to the hexadirectional signal if the preferred directions of these cells align with the axes of the cells' grid pattern (Doeller et al., 2010). Technological advances such as two-photon microscopy (Dombeck et al., 2007) enabling imaging of cellular responses in larger portions of the rodent brain might allow future research to shed further light on the dynamics underlying hexadirectional signals on the population level.

The firing of place and grid cells conveys positional information to navigate Euclidean space. Here, we hypothesize that the spatially constrained firing of place cells and the metric provided by the entorhinal grid system might provide a domain-general mechanism to map dimensions of experience. In this framework, the activity of place cells can be conceived as indexing locations in a cognitive space spanned by the entorhinal grid system. Three further neural coding mechanisms identified in spatial navigation studies illustrate how the hippocampal-entorhinal system may support a core mechanism of mapping cognitive spaces. First, the firing

fields of place and grid cells increase in size along the dorsoventral axis of the rodent hippocampus (Hafting et al., 2005; Jung et al., 1994; Barry et al., 2007; Kjelstrup et al., 2008; Stensola et al., 2012), in line with mapping of cognitive spaces at different levels of granularity for multi-scale representations of knowledge hierarchies or nested conceptual information. Second, the ability of place cells to undergo global remapping (Bostock et al., 1991; Leutgeb et al., 2004; Muller and Kubie, 1987) allows the flexible formation of a multitude of uncorrelated maps for different cognitive spaces, which can be reinstated via attractor dynamics (Marr, 1971; McNaughton and Morris, 1987; Hopfield, 1982; McClelland et al., 1995). Third, sequential activity of place (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Foster, 2017) and grid cells during replay (Ólafsdóttir et al., 2016; Gardner et al., 2017; Trettel et al., 2017) and theta oscillations (Johnson and Redish, 2007) enables the simulation of trajectories (Buckner, 2010; Byrne et al., 2007; Hasselmo, 2011) through different locations in a cognitive space for adaptive cognition and behavior.

A CONTINUOUS MAP OF EXPERIENCE

The spatially constrained firing of place and grid cells provides a continuous code for the dimensions of space, in which neighboring positions have similar representations due to partially overlapping firing fields across the population of cells (Fig. 2AB). We build on findings that the continuous code of spatially selective cells maps dimensions of experience beyond Euclidean space, which affords flexible cognition via the formation of cognitive spaces, whose dimensions are geometrically constrained as described above.

Contrary to rodents, the visual system has evolved to be the dominant source of sensory information in primates. During visual exploration of naturalistic images, neurons in primate EC encode gaze position with six-fold symmetric firing patterns that are the hallmark of grid cell firing during navigation (Killian et al., 2012). Further, some entorhinal neurons preferentially discharged when monkeys explored the edges of the visual stimuli (Killian et al., 2012), akin to the firing of border cells in rodents (Lever et al., 2009; Savelli et al., 2008; Solstad et al., 2008). Others were tuned to saccade direction (Killian et al., 2015), putatively signaling similar information as head direction cells (Cullen and Taube, 2017). Together with two recent reports of eye-movement dependent hexadirectional signals in human EC (Nau et al., 2018a; Julian et al., 2018) these findings support the notion

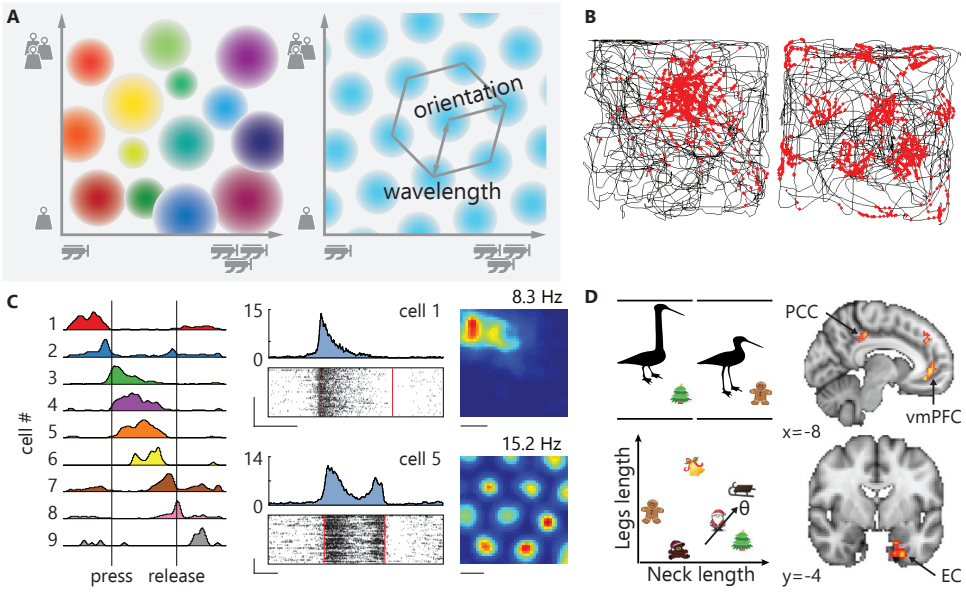


Figure 2 | Place and grid cells map dimensions of cognitive spaces.

(**A**) Left: Colored circles illustrate firing fields of hippocampal place cells in a space spanned by the dimensions of car engine power and weight. Each color represents one individual place cell. Collectively, the firing fields map all locations of the space. Right: The hexagonally symmetric firing pattern of grid cells (one cell's pattern is shown) provides a metric for the dimensions of the space. (**B**) Firing of a place cell (left) and a grid cell (right) recorded from the hippocampus and entorhinal cortex, respectively. Spike locations (red dots) are shown on the animal's path (black line) through a square enclosure (adapted from Moser et al., Annual Review of Neuroscience (2008)). (**C**) Frequency fields of place and grid cells. Left: While the animal presses a lever to manipulate the frequency of a tone, different cells fire at different frequencies during the frequency sweep. Center and right: Two example cells active in the sound-modulation task identified as a place cell and a grid cell, respectively, during navigation. Center panels show peristimulus histograms with firing rate in hertz (top) and raster plots displaying spikes as a function of sound frequency for different trials (bottom). Right panels show spatial firing rate maps with maximum firing rate (adapted from Aronov et al., Nature (2017)). (**D**) Hexadirectional signals in two-dimensional bird space. Left: Participants learned associations of Christmas symbols and bird silhouettes with variable leg and neck length. Arrow illustrates example trajectory with angle q. Right: During trajectories through this space defined by morphing birds, hexadirectional signals were observed in the entorhinal cortex (EC) and a network of brain regions implicated in mnemonic processing (PCC, posterior cingulate cortex; vmPFC, ventromedial prefrontal cortex, adapted from Constantinescu et al., Science (2016)).

that representations of visual space follow coding principles identified in spatial navigation research in rodents, suggesting a shared neural basis.

Next to its role in parsing sensory information, this spatial mapping mechanism might also encode a dimension inherent to all experience: time. During space-clamped running throughout a temporal delay, so-called time cells preferentially fire at specific time points (Pastalkova et al., 2008; MacDonald et al., 2011). The populations of these time cells in the hippocampus and EC overlap substantially

with the population of place and grid cells, respectively, suggesting that cells in these regions might exhibit mixed selectivity for space and time (Kraus et al., 2013, 2015). While these hippocampal time cells map repeated intervals in the range of seconds, it remains elusive how time might be discriminated at larger scales. For this, an ongoing temporal signal slowly drifting across hours and days might be transmitted to the hippocampus from the lateral EC, enabling the tagging of time and order to events for mnemonic processing (Tsao et al., in press).

The mapping of dimensions by the hippocampal-entorhinal system might be a general mechanism to code for task-relevant information. In a sound modulation task, rats manipulated the frequency of an auditory stimulus to 'navigate' to a target frequency (Aronov et al., 2017). Cells in the hippocampus and EC exhibited discrete firing fields with elevated firing constrained to each cell's preferred frequency range, thereby mapping the continuous dimension of sound frequency so that neighboring frequencies exhibited similar firing patterns across cells (Fig. 2C). A subset of the cells carrying frequency-specific information in the sound modulation task could be characterized as place and grid cells during random foraging (Aronov et al., 2017). Hippocampal representations of continuously changing odor concentrations suggest that the directional selectivity of place cells on a linear track might be preserved when mapping a non-spatial dimension (Radvansky and Dombeck, 2018).This suggests a flexible recruitment of cells to map task-relevant dimensions of experience while maintaining spatial coding properties.

However, we are constantly faced with complex stimuli defined by not one but multiple feature dimensions. Spatial coding has been demonstrated for stimuli varying along two independent dimensions. In one recent study, human participants learned to associate target objects with pictures of stick-figure birds (Fig. 2D) differing in the length of their necks and legs (Constantinescu et al., 2016). While undergoing fMRI, participants watched trajectories through this two-dimensional feature space and BOLD-signals were analyzed as a function of the angular orientation of these trajectories following the analysis developed to investigate grid coding during virtual navigation (see Box; Doeller et al. (2010)). Hexadirectional signals were observed in EC as well as a network of regions (Constantinescu et al., 2016) involved in human autobiographical memory (Byrne et al., 2007; Schacter et al., 2007; Ranganath and Ritchey, 2012) and exhibiting hexadirectional signals during navigation (Doeller et al., 2010). Within participants, the orientation of the hexadirectional signal was consistent across frontal and medial temporal brain areas and stable over separate sessions more than a week apart (Constantinescu et

al., 2016), suggesting a role of the entorhinal grid system in storage and retrieval of consolidated conceptual knowledge.

In our framework, place and grid cells not only map Euclidean space for navigation, but also map cognitive spaces spanned by relevant feature dimensions (Fig. 2A). Building upon the geometric description of cognitive spaces, this mapping provides an account for how multi-dimensional spaces of experience can be instantiated in the hippocampal formation. If stimuli are located in the space based on their characteristic values on the feature axes, place cell firing might encode stimuli at specific locations in cognitive space based on the respective values along the feature dimensions (Gärdenfors, 2000). Similar stimuli are located at nearby positions, whereas dissimilar stimuli might be located further apart in cognitive space. Distances between positions are reflected in population vectors of place cell activity, which will be more similar for nearby positions due to overlapping firing fields of cell ensembles. Therefore, the mapping of cognitive spaces by cell populations in the hippocampal-entorhinal region provides a mechanism to generate similarity between stimuli (Gärdenfors, 2000), a concept central to generalization and planning (Shepard, 1987). Applying this to the introductory example of a two-dimensional space of engine power and weight, positions encoded by constrained firing of place cells reflect specific combinations of values along the dimensions, for example high engine power and low weight. This might fall in a region representing the concept of racing cars, which would typically be characterized by this feature combination. The convexity of conceptual regions enables the generalization that a car whose engine power and weight fall between those of two known sports cars also belongs to this category (Gärdenfors, 2000). In this framework, the regular firing patterns of grid cells might provide a metric (Bush et al., 2015; Hafting et al., 2005; Herz et al., 2017; McNaughton et al., 2006) for the dimensions of cognitive spaces. This would allow not only for the encoding of positions but also their relations in a similar fashion as proposed for spatial navigation and memory (Buzsáki and Moser, 2013; Moser et al., 2017). Thereby, the grid system might provide the coordinate system for the dimensions of cognitive spaces, in line with its suggested role in extracting dimensions from task states (Dordek et al., 2016; Stachenfeld et al., 2017). However, the precise nature of the metric remains elusive as different metrics can underlie cognitive spaces (Gärdenfors, 2000) and the impact of deviations from regular hexagonal firing patterns (Barry et al., 2007; Krupic et al., 2015; Stensola et al., 2015) is unclear.

Neural recordings in patients suggest the existence of cells in the hippocampus and EC that selectively respond to a narrow range of stimuli (Quiroga et al., 2005).

Despite the limitations imposed by the necessity of stimulus selection, the claim that this phenomenon reflects coding of conceptual entities such as famous people or places is supported by the preservation of responses across stimulus modalities. Even though there are no experimentally defined feature axes in these studies, these cells sometimes respond similarly to entities that appear nearby in cognitive space. For example, a cell responding to pictures of the Tower of Pisa also exhibited increased firing in response to the Eiffel Tower (Quiroga et al., 2005). This is in line with the assumption that proximal positions in cognitive space are represented by overlapping cell assemblies, comparable to the representations of nearby locations in space by overlapping place cell populations.

The actions and positions of other agents in space are central to interacting with and learning from conspecifics (Stolk et al., 2016). In bats and rodents, hippocampal place cells of an observer animal encode the location of a conspecific or a moving object, indicating the tracking of relevant agents in the environment (Danjo et al., 2018; Omer et al., 2018). The human hippocampus also signals the position of others along interpersonal dimensions. In one study, participants were asked to respond to statements of fictitious characters located in a two-dimensional space of power and affiliation (Tavares et al., 2015). Hippocampal activity tracked the position of the counterpart in this social space at times of interaction, exhibiting greater activity during interaction with counterparts with higher power and affiliation (Tavares et al., 2015). Hippocampal encoding of the power dimension has also been demonstrated after learning of social hierarchies (Kumaran et al., 2016a). These results suggest a hippocampal involvement in the representation of others in both navigable and interpersonal spaces.

It is thus conceivable that coding of fundamental dimensions of experience also underlies the involvement of the hippocampal-entorhinal system in episodic memory, where individual episodes are considered to be embedded in a spatio-temporal context (Schiller et al., 2015). As place and grid cells carry information about both space and time they might thereby signal the context in which events are experienced. Evidence for this stems from a study in which participants encountered objects along a fixed route through a virtual city. Through learning, the similarity structure of multi-voxel activity patterns in the hippocampus changed to reflect remembered spatial and temporal distances between object pairs (Deuker et al., 2016). This is in line with the notion that the hippocampus encodes the dimensions of space and time along which episodic experiences are organized and the central role of the hippocampal formation in binding stimuli or events to a context in service

of episodic memory (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 2014; Ekstrom and Ranganath, 2017).

Together, these findings suggest that the firing of functionally defined cell types in the hippocampal-entorhinal system prevails across task-relevant dimensions to map dimensions of experience in cognitive spaces. Stimuli are arranged in a spatial format where similarity between positions is reflected in the distance along the dimensions spanning the cognitive space. The representation of cognitive spaces allows not only associative or transitive inference, accounted for via overlapping relational networks in the realm of relational memory theory (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 1988, 2014), but also generalization and inference to novel stimuli and situations. Below we describe how key coding principles of the hippocampal-entorhinal system make it an ideal candidate for entertaining cognitive spaces.

MULTIPLE SCALES OF CODING

One hallmark of abstract knowledge is the representation of information at different hierarchical levels (Collins and Quillian, 1969; Tenenbaum et al., 2011). For example, you might identify a particular vehicle characterized by low weight and high engine power as a Porsche (Fig. 3A). On a more general level, you might classify it as a sports car allowing you to infer some properties such as high driving speed, whereas on a more specific level you might wonder about the particular model and its associated characteristics. A likely mechanism for learning and representing information at different scales (Collins and Quillian, 1969) is described in studies investigating the response properties of place and grid cells along the dorsoventral axis of the rodent hippocampal formation (Fig. 3B) (Hafting et al., 2005; Jung et al., 1994; Barry et al., 2007; Kjelstrup et al., 2008; Stensola et al., 2012). On an 18m long linear track, the width of hippocampal place cell firing fields increased from less than 1m in the dorsal to approximately 10m in the ventral hippocampus (Kjelstrup et al., 2008). In human fMRI, this might be reflected in a voxel-similarity gradient along the hippocampal long-axis (Brunec et al., 2018a). Similarly, the scale of entorhinal grid cells increases from the dorsal to the ventral parts of the medial EC, reflecting larger firing fields and larger spacing between fields (Hafting et al., 2005; Barry et al., 2007; Stensola et al., 2012; Brun et al., 2008). In contrast to place cell firing field width (Kjelstrup et al., 2008), grid scale changes in discrete steps between modules of cells sharing a similar scale and orientation (Hafting et al., 2005; Barry et al., 2007; Stensola et al., 2012; Krupic et al., 2015).

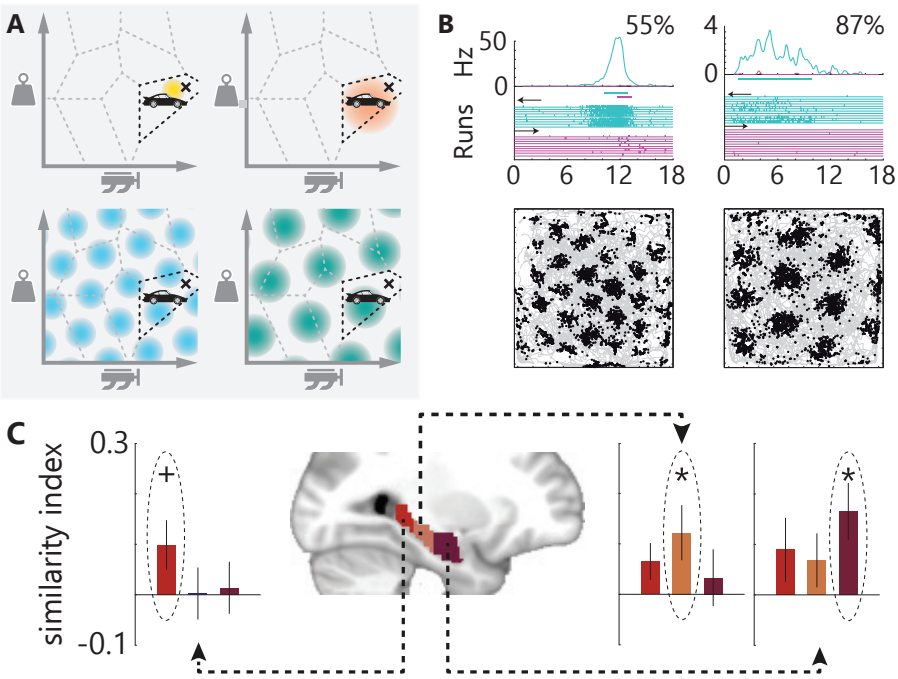


Figure 3 | Multiple levels of representation.

(**A**) Representations of information at different levels of granularity can be supported by multiple spatial scales in the hippocampal formation. Narrower (left column) or broader (right column) positions in cognitive spaces can be encoded by place cells (top row) and grid cells (bottom row), respectively. Schematics illustrate firing fields of four different cells. Higher-level information (e.g., about the concept of sports cars) can be ascribed to a lower-level stimulus (black cross). For example, when learning that a specific car model is a sports car, we can infer that it is likely to have high engine power. (**B**) The firing field size of place cells (top) and the size and spacing of grid firing fields (bottom) increase from dorsal to ventral recording sites in the rodent hippocampal-entorhinal system. Top panels show firing rates in hertz as a function of position along the linear track. Percentages indicate location along the dorsoventral axis. Bottom panels show spike locations (black) overlaid on the animal's path (gray) (adapted from Kjelstrup et al., Science (2008) and Stensola et al., Nature (2012)). (**C**) The granularity of mnemonic networks scales along the long axis of the human hippocampus with pairwise associations of elements in posterior and integrated networks in anterior hippocampus. Bars show model evidence (mean ± SEM) for mnemonic networks of small (red), medium (orange), and large (purple) scale. *P ≤ 0.05, +P = 0.053 (adapted from Collin et al., Nature Neuroscience (2015)).

Different scales of information represented at distinct anatomical locations of the hippocampal formation might serve as a general mechanism across different stimulus domains. Positional decoding in cognitive spaces might benefit from the combination of multiple scales of information analog to navigable space (Bush et al., 2015; Herz et al., 2017). The spacing of firing fields of grid cells responding to locations in visual space increases with anatomical distance to the rhinal sulcus (Killian et al., 2012), which approximately reflects the anterior-posterior axis of the primate hippocampus

corresponding to the dorsal-ventral axes of the rodent hippocampus (Strange et al., 2014). This is further paralled by broader tuning of saccade direction cells as anatomical distance to the rhinal sulcus increases (Killian et al., 2015).

Multiple scales of coding along the dorsoventral axis of the rodent hippocampus are in line with the scale increase of mnemonic networks represented along the anterior-posterior axis of the human hippocampus (Fig. 3C) in participants watching videos of life-like events forming narratives (Collin et al., 2015; Milivojevic and Doeller, 2013). The scale at which these mnemonic networks were represented differed across the hippocampus with the posterior portion representing associations of the most recently linked pair of events, while the mid-portion held information about multiple event pairs and patterns in the anterior hippocampus were indicative of integrated networks of all events in a narrative (Collin et al., 2015). Multi-scale event representations, interacting with mnemonic processing in the hippocampus, have been dissociated along the cortical hierarchy in humans (Baldassano et al., 2017). Representing cognitive spaces at different scales allows for the generalization of specific experience and the formation of contextual codes via more global representations. In rodents performing a context-dependent object discrimination task, ventral hippocampal neurons exhibited responses generalizing across events within a spatial context while strongly distinguishing between contexts after extended learning, whereas dorsal hippocampal neurons discriminated events within the same context with activity patterns reflecting the hierarchical task structure (Komorowski et al., 2013; McKenzie et al., 2014). The representation of integrated codes for overlapping memories might further enable inferential reasoning about related memories. In humans, transitive inference and generalization are supported by anterior portions of the hippocampus (Preston et al., 2004), whereas more posterior portions are associated with the retention of original memories and element segregation when associations overlap (Kuhl et al., 2010; Schlichting and Preston, 2015). These findings suggest a mnemonic gradient along the hippocampal long-axis in humans paralleling the differences in granularity of spatial representations along the dorsoventral axis of the rodent hippocampus. The most detailed representation might allow for fine-grained discrimination of locations in cognitive space, whereas the representation of larger areas might enable inference and the generalization of behavior to never-experienced stimuli and situations, not limited by the need for associations between nodes in relational mnemonic networks (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 1988, 2014).

Representing knowledge at different levels of granularity in cognitive spaces

requires a cognitive code signaling positions in this space at different resolutions. Generalizing from the above findings to conceptual information, it appears plausible that the gradient of granularity plays an important role in learning and representing hierarchical knowledge structures (Collins and Quillian, 1969; Tenenbaum et al., 2011). In these, overarching categories can be conceived as larger areas of cognitive space, putatively represented by place and grid cells with larger firing fields. Subcategories at a finer scale would then correspond to locations nested within these larger areas. Such a nested representation of cognitive spaces could enable inference via the transfer of knowledge from the superordinate category to new exemplars or subcategories (Gärdenfors, 2000).

FLEXIBLE FORMATION OF STABLE COGNITIVE SPACES VIA REMAPPING AND ATTRACTOR DYNAMICS

The hippocampus has been shown to contribute to a variety of cognitive domains (Hannula and Duff, 2017). If the hippocampal-entorhinal system maps multitudes of cognitive spaces, this system needs to exert remarkable flexibility not only in terms of the dimensions it can represent, but also to rapidly switch between cognitive spaces (Fig. 4A). This flexibility is demonstrated by the capacity of hippocampal place cells to undergo global remapping (Bostock et al., 1991; Leutgeb et al., 2004; Muller and Kubie, 1987). Different subsets of hippocampal cells will exhibit place fields in two different environments with spatial relationships among cells active in both environments not being maintained, rendering the two maps orthogonal to each other (Fig. 4B) (Leutgeb et al., 2004). In contrast to place cells, entorhinal grid patterns have varying offsets to boundaries in different environments, but maintain their spatial phase relative to each other, resulting in consistent relations between the firing fields of grid cells (Fyhn et al., 2007; Yoon et al., 2013). These relationships are maintained not only between environmental contexts but also across behavioral states, with essentially identical cross-correlation patterns exhibited by populations of grid cells and other medial entorhinal cells during free foraging and slow-wave or REM sleep (Gardner et al., 2017; Trettel et al., 2017).

Remapping-like behavior of hippocampal cells has also been observed for time cells encoding temporal intervals during the delay of a memory-based discrimination task (MacDonald et al., 2011). When the length of the delay was altered, a subset of time cells remapped. Some cells ceased to be active, became active at different times of the delay or previously silent cells became active, whereas other cells maintained

their firing during absolute or relative times of the altered delay (MacDonald et al., 2011). The flexible recruitment of cells to map positions in cognitive spaces is further illustrated by hippocampal representations of conspecifics, where subsets of place cells exclusively encoded the location of the self, the conspecific or an inanimate object, whereas other cells exhibited firing fields for both the self and other, but at different locations (Omer et al., 2018). The decorrelation of hippocampal representations has also been observed using multivariate pattern analyses of human episodic memory. Reconfigurations of hippocampal activity patterns reflect associations and narrative insight in increased similarity, but also reduce overlap resulting in decorrelation through experience (Milivojevic et al., 2015; Schapiro et al., 2012).

If different cognitive spaces are represented by orthogonal subsets of hippocampal cells, how can already-formed representations of spaces be reinstated to provide stable maps over multiple encounters? In rodents, reinstatement of place cell firing patterns has been observed on a trial-by-trial level indicating rapid switching between maps upon re-exposure to a highly familiar circular and square enclosure (Figure 4C), respectively (Wills et al., 2005). A likely candidate mechanism governing the reactivation of established cognitive spaces may involve attractor networks (Marr, 1971; McNaughton and Morris, 1987; Hopfield, 1982; McClelland et al., 1995). In intermediate steps of a morph sequence transforming a square to a circular environment, hippocampal patterns resemble the original map until a switch point at which the representation is pulled towards exhibiting the other map (Wills et al., 2005), though progressive transformations of hippocampal codes have been observed under different experimental conditions (Leutgeb et al., 2005). Neuroimaging research points towards a role for stored hippocampal representations in perceptual discrimination (Bonnici et al., 2012) and attractor dynamics specifically in memory-guided human behavior (Fig. 4D) (Steemers et al., 2016). Participants learned positions of identical sets of objects in two virtual environments distinguished by background cues and were subsequently tested in a series of morph environments following a linear transformation between the original environments. Spatial memory responses in the intermediate environments followed the sinusoidal pattern predicted by the influence of attractor networks and were paralleled by corresponding non-linear changes of hippocampal activity patterns (Steemers et al., 2016).

These findings indicate that response properties of the hippocampal-entorhinal system enable the formation of independent maps for distinct cognitive spaces to map

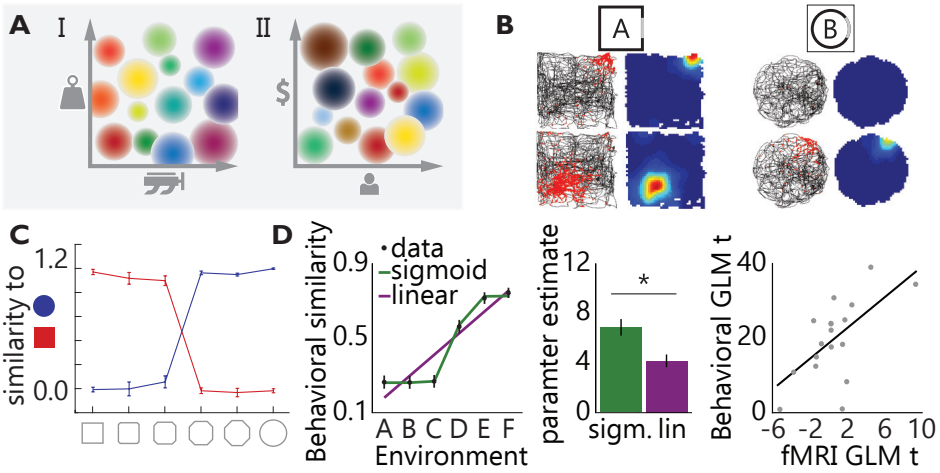


Figure 4 | Remapping and attractor dynamics for flexible cognitive spaces.
(**A**) Different task-relevant dimensions (e.g., a car's passenger capacity and price instead of its engine power and weight) can be mapped by the recruitment of a different subset of place cells and the re-arrangement of firing fields between spaces. (**B**) In navigating rodents, place cells remap to represent different environments. Spike locations (red) overlaid on the animal's path (black) and rate maps (warmer colors indicating increased firing) are shown for two place cells (rows) in two environments (columns). One place cell is active in the square environment but not in the circular environment, whereas the other exhibits a firing field in both environments but at unrelated positions (adapted from Fyhn et al., Nature (2007)). (**C**) Attractor dynamics enable rapid switches between established maps. The similarity to established maps of a square environment (red) and a circular environment (blue) is shown for a sequence of intermediate environments. The data show a sigmoidal rather than linear shift function. (**D**) Attractor dynamics in human spatial memory. After learning object locations in two base environments, A and F, participants were tested in a sequence of intermediate environments. Left: Spatial memory responses exhibited patterns better explained by a sigmoidal function than by a linear function. Center: Hippocampal multi-voxel pattern similarity was better predicted by a sigmoidal model than by a linear model of behavior. Right: The fit of a canonical sigmoidal function to spatial memory responses was associated with hippocampal pattern similarity values across participants. T-values from general linear modeling (GLM) reflect fit of sigmoidal model to hippocampal fMRI and behavioral data. Error bars in (C) and (D) reflect SEM. *P < 0.05 (adapted from Steemers et al., Current Biology (2016)).

different dimensions. Consider again the car example: In a different context, weight and engine power might not be the most relevant feature dimensions, but rather a new space might be spanned by the dimensions of possible number of passengers and price (Fig. 4A). Via hippocampal remapping, positions in the new space can be mapped and the positions of different cars are re-defined based on their feature values on the new dimensions, resulting in new similarity relations. The more rigid firing properties of the entorhinal grid system might provide a stable metric through its intrinsic coherence across spaces. Once established, rapid switching between the maps of different cognitive spaces through remapping demonstrates the flexibility of the system and permits the representation of relevant information based on behavioral context. Switches between maps might be triggered by internal or external cues,

attentional shifts or changes in task demands rendering other dimensions relevant. Attractor dynamics might serve as anchors for stable representations, which enable learning over multiple encounters as well as the generalization of behavior across similar experiences.

SIMULATIONS AND READ-OUT OF TRAJECTORIES FOR DECISIONS

Recording studies in rodents have revealed structured activity of hippocampal place cells during rest in which cells fire sequentially, resulting in trajectories reflecting past experience on a maze (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996). This so-called 'replay' demonstrates the ability of spatially tuned cells to represent locations beyond the animals current position (Foster, 2017). During replay sequences, place cells maintain their spatial relationships to each other (Wilson and McNaughton, 1994; Diba and Buzsáki, 2007), resulting in trajectories through space occurring both in forward and reverse order (Diba and Buzsáki, 2007; Davidson et al., 2009). Place cell sequences during replay are temporally compressed compared to trajectories during running (Skaggs and McNaughton, 1996; Foster, 2017; Lee and Wilson, 2002) and might thereby allow fast simulations of trajectories through cognitive space (Fig. 5AB). Sequences replaying trajectories in reverse order might evaluate previous paths and associate the states visited with reward information for learning of adaptive behavior (Foster and Wilson, 2006; Gupta et al., 2010; Dupret et al., 2010; Foster, 2017). Hippocampal sequences may extend to the simulation of potential future paths and their outcomes. In a spatial alternation task, replay of both correct and incorrect future trajectories supports learning and planning (Singer et al., 2013) and the disruption of sequences during sharp-wave ripples impairs successful performance (Jadhav et al., 2012). Further, place cell sequences reflect future paths during goal-directed behavior (Pfeiffer and Foster, 2013) and trajectories extending into the shock zone of a linear track prior to avoidance behavior (Wu et al., 2017). These findings dovetail with extensive research on place cell sequences during ongoing theta oscillations, where trajectories towards potential goals are represented within different theta cycles (Johnson and Redish, 2007), suggesting simulations of possible trajectories through cognitive space.

What role do mental simulations of trajectories play for planning and decision-making in humans? During navigational planning on a circular track, hippocampal activity patterns carry information not only about the start and goal location,

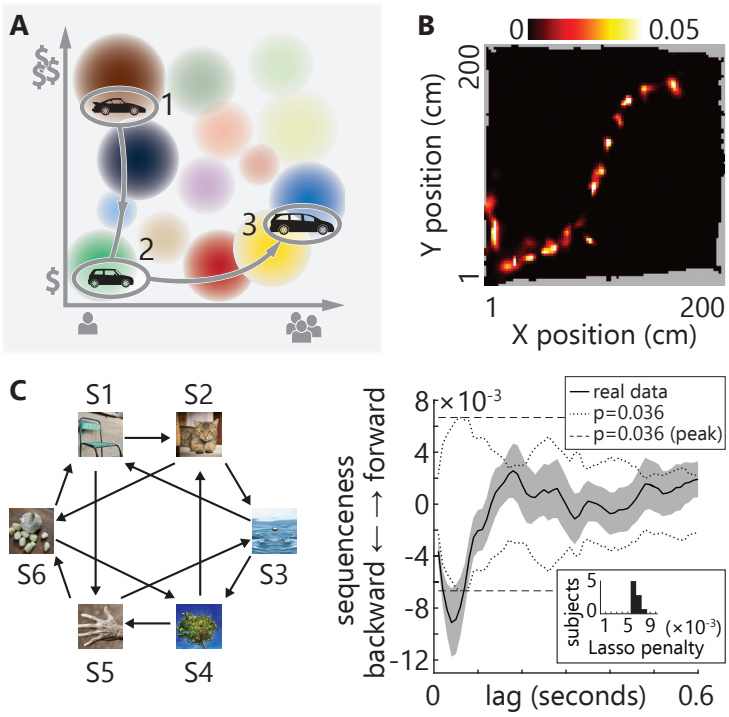


Figure 5 | Hippocampal simulations via sequential activity.

(**A**) Sequential activity simulates different positions (1 to 3) in cognitive space, allowing the evaluation of different car types along the trajectory when deciding which car to buy. (**B**) Simulation of trajectory in a square enclosure. The posterior probabilities of positions from Bayesian decoding applied to different timeframes during a sharp wave ripple event result in a trajectory through space (adapted from Pfeiffer & Foster, Science (2015)). (**C**) Reverse trajectories through state space. Left: States are represented by images; arrows indicate possible transitions between states. Right: The solid black line shows a "sequenceness" measure indicating the probability of decoding a successive or preceding state (positive and negative values, respectively) at different time points after the decoding of a state during planning. Reverse sequences were observed with a lag of around 40 ms between state space positions. Inset shows histogram of lasso penalties from logistic regression (adapted from Kurth-Nelson et al., Neuron (2016)).

but also locations along the optimal rather than suboptimal path, in accordance with sequential simulations of trajectories through space (Brown et al., 2016). The entorhinal grid system has also been linked to imagining navigation (Horner et al., 2016) and snapshots from stationary viewpoints (Bellmund et al., 2016), implicating it in the planning of trajectories through space in line with replay in the EC (Ólafsdóttir et al., 2016; Gardner et al., 2017; Trettel et al., 2017; O'Neill et al., 2017) and the observation of grid cell firing during covert attentional trajectories through visual space (Wilming et al., 2018). Beyond navigable space, the hippocampal-entorhinal system extracts statistical regularities of non-spatial

sequences (Kumaran and Maguire, 2006; Hsieh et al., 2014) and forms relational maps of the underlying structures the sequences were derived from (Schapiro et al., 2012; Garvert et al., 2017), potentially drawn upon to plan trajectories through task spaces. Indeed, in a non-spatial decision-making task, sequential reactivations of previously visited states reflected reverse trajectories (Fig. 5C) through a space of discrete states represented by objects (Kurth-Nelson et al., 2016). While the nature of the magnetoencephalography (MEG) signal and the analysis approach focusing on visual responses make a direct hippocampal origin of these results unlikely, they might reflect sequential reinstatement of visual representations orchestrated by the hippocampus (Kurth-Nelson et al., 2016). Initial evidence suggests sequential hippocampal activity can be observed using fMRI (Schuck and Niv, 2018). Episodic cues can elicit memory-guided simulation of past experience influencing choice behavior (Bornstein and Norman, 2017) and prospective simulations of trajectories through a task's state space have been linked to model-based behavior in a two-stage decision-making task (Doll et al., 2015). These findings indicate that trajectories through cognitive spaces representing task states can be simulated based on prior experience.

Cognitive spaces enable generalization and finding novel trajectories via the representation of positions along defined dimensions. The rodent hippocampus can recombine separate trajectories across segments of a two-choice T-maze, usually not experienced successively, to infrequently construct never-experienced place-cell sequences (Gupta et al., 2010). Similarly, hippocampal place cells constructed trajectories through previously non-traversed space after rats observed the placement of a reward on one T-maze arm while confined to the stem of the maze (Ólafsdóttir et al., 2015). In humans, hippocampal simulations have been linked to the construction of imaginary scenarios (Hasselmo, 2011; Buckner, 2010; Byrne et al., 2007; Schacter et al., 2007), an ability impaired in patients with hippocampal lesions (Hassabis et al., 2007a). Mental simulations can be conceived as putative trajectories through the space of episodic experience (Byrne et al., 2007; Hasselmo, 2011) and have been shown to influence decision-making in a delay-discounting task (Peters and Büchel, 2010). Furthermore, in a task in which participants imagined novel compound goods consisting of two familiar foods the hippocampus and mPFC flexibly combine past experiences (Barron et al., 2013). These findings show that the hippocampus can flexibly draw on past experience to form and simulate novel trajectories through cognitive space allowing adaptive decision-making and behavior.

Simulations of trajectories building on spatially tuned cells have been incorporated in models of episodic memory and consolidation (Byrne et al., 2007; Hasselmo, 2011; Kumaran et al., 2016b). Replay-inspired simulation of experience has also been used to integrate reinforcement learning and deep neural networks in artificial intelligence (Mnih et al., 2015). Here, we suggest that contemplating a number of stimuli can be conceived as a trajectory through cognitive space. In line with its proposed role in future anticipation and prediction (Buckner, 2010; Byrne et al., 2007; Schacter et al., 2007; Stachenfeld et al., 2017; Hassabis et al., 2007a), the hippocampal-entorhinal system supports these trajectories via sequential activity of spatially tuned cells. The entorhinal grid system might span up a space based on a set of dimensions and thereby provide the framework for flexible simulations of positions and trajectories by the hippocampus. Relevant feature dimensions can range from locations in a maze to abstract state spaces in decision-making problems. Novel trajectories can be generated from past experience and trajectories can be simulated via sequential hippocampal activity to guide future behavior. Building on the geometric characterization of cognitive spaces, experience can be generalized to the outcome of novel trajectories or actions. For example, if of two prior actions one under- and one overshot a goal, then an intermediate trajectory through cognitive space will approximate the goal more closely. While it has been suggested that sequential activity might reflect inherent hippocampal dynamics (Dragoi and Tonegawa, 2011; Grosmark and Buzsáki, 2016), other accounts highlight hippocampal interactions with state-space representations in frontal regions in sequence generation and action selection, and the role of striatal regions in sequence evaluation (Peters and Büchel, 2010; van der Meer et al., 2010; Schuck et al., 2016; Wikenheiser and Schoenbaum, 2016; Kaplan et al., 2017a).

OPEN QUESTIONS AND FUTURE DIRECTIONS

We used spatial navigation investigated in rodents and humans (Moser et al., 2017; Epstein et al., 2017) as a model system to identify key neural mechanisms and fused them with concepts from cognitive science to describe the central neural coding machinery underlying higher-level cognition in humans. Future research should help to elucidate the generation of cognitive spaces and their governing principles. For example, the conditions of the generation of the continuous code in the hippocampal-entorhinal system are still largely unclear. This concerns both the circuit level interactions of brain regions (Moser et al., 2014) and their maturation (Donato et al., 2017), the role of spatial cognition during development (Hermer and

Spelke, 1994) and the potential breakdown of spatial codes in aging and disease (Kunz et al., 2015; Stangl et al., 2018; Fu et al., 2017). Of particular interest is how precisely the entorhinal grid code emerges in its hexagonally symmetric form. Advances in techniques such as two-photon calcium imaging (Dombeck et al., 2007) will foster the population level understanding of neural codes in rodents. In concert with the rise of high-resolution fMRI and optimized MEG protocols, this should further bridge the gap between systems and cognitive neuroscience to unravel neural mechanisms promoting refinements of processing in cognitive spaces; and spur the mapping of brain structures to specific functions.

Cognitive spaces can be multi-dimensional. Therefore, the question how a continuous code can be extended to map out additional dimensions is fundamental. Research in humans (Kim et al., 2017), rats (Hayman et al., 2011) and bats has shed light on how spatial coding of place (Yartsev and Ulanovsky, 2013) and head direction cells (Finkelstein et al., 2015) can be extended to the third dimension. However, grid-like coding in 3D as well as evidence for spatial coding in cognitive spaces of higher dimensionality remains elusive. Related to this is the question how cognitive spaces are spanned by multiple, potentially integral, dimensions. Further, the nature of the metric underlying cognitive spaces can be investigated using measures of neural and behavioral similarity between positional representations. While the assumption of a Euclidean metric (Kosslyn et al., 1978) might be most intuitive when comparing cognitive to navigable space, there is evidence for topological representations of spaces in both rodents (Dabaghian et al., 2014) and humans (Garvert et al., 2017; Pylyshyn, 1973; Warren et al., 2017).

While grid cells have been implicated in representing dimensions of cognitive spaces, the extent to which they retain their specific firing patterns remains to be explored. For instance, is the modular organization of the EC stable across cognitive spaces? One might expect that grid cells from the same module also show similar orientation and spacing when representing dimensions of a given cognitive space, but differ in their spatial phase from the boundaries of the space. Generalizing from this question, the role of other cell types encoding spatial information in the context of navigation remains to be elucidated. For example, one could envision a role for border cells in signaling event (Ben-Yakov et al., 2013; Brunec et al., 2018b; Ezzyat and Davachi, 2014) or conceptual boundaries when learning to categorize stimuli drawn from a cognitive space as belonging to different concepts. Likewise, head direction (Cullen and Taube, 2017), goal direction (Sarel et al., 2017) or object vector (Høydal et al., 2018) cells might be involved in representing relationships

between stimuli located at different locations in a cognitive space. Similarly, the role networks of brain regions acting in concert with the hippocampus (Byrne et al., 2007; Constantinescu et al., 2016; Ranganath and Ritchey, 2012; Schacter et al., 2007) play for cognitive spaces and differences to the encoding of multidimensional stimulus spaces in other brain regions, for example during face processing in monkey inferotemporal cortex (Chang and Tsao, 2017), should be explored.

An intriguing question concerns the measurement of behavioral benefits of cognitive space formation, for example via the generation of shortcuts through cognitive space or the impaired ability to find shortcuts in lesion patients. Further, how does information encoded in different cognitive spaces interact? Can trajectories encoded in one space be transferred to another and be retrieved there to guide behavior? Contrarily, it might be possible to bring codes from different cognitive spaces into conflict with one another, which might result in interference across spaces, or investigate effects of deformations of firing patterns (Barry et al., 2007; Krupic et al., 2015; Stensola et al., 2015) across spaces to further elucidate how representations of different spaces are entertained by overlapping neural substrates.

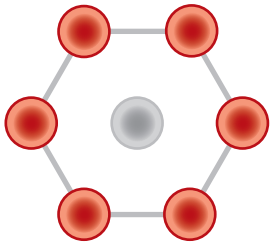
CONCLUSION

In this theoretical article, we propose cognitive spaces as a primary representational format for information processing in the brain. Combining key mechanisms identified in systems neuroscience and concepts from cognitive science and philosophy, we developed a cognitive neuroscience framework for processing and representing information in cognitive spaces in the hippocampal-entorhinal system. Place and grid cells might have evolved to represent not only navigable space, but to also map dimensions of experience spanning cognitive spaces governed by geometric principles. In these cognitive spaces, stimuli can be located based on their values along the feature dimensions mapped by place and grid cells. These spatially specific cells provide a continuous code that allows similar stimuli to occupy neighboring positions in cognitive space, encoded by overlapping population responses. In this framework, concepts are represented by convex regions of similar stimuli. The multi-scale spatial code along the long-axis of the hippocampal formation enables representing stimuli at different granularities for both generalization and maintenance of fine details in hierarchical knowledge structures. Ever-changing demands requiring the flexible mapping of different dimensions of relevance are met by the capacity of the hippocampus to remap to flexibly form cognitive spaces for

which the low-dimensional entorhinal grid code might provide a stable metric. An established mapping of a cognitive space might be reinstated via attractor dynamics and pattern completion to provide stable representations of familiar dimensions. Experiencing a sequence of stimuli results in a trajectory through cognitive space. We propose that sequential hippocampal activity in the form of replay and theta sequences allows simulations of temporally compressed trajectories through cognitive spaces for flexible cognition and adaptive behavior. In sum, we suggest cognitive spaces as a domain-general format for human thinking, thus providing an overarching framework, which can also help to better understand cognitive breakdown in neurodegenerative diseases (Fu et al., 2017; Kunz et al., 2015) and to inform novel architectures in artificial intelligence (Mnih et al., 2015).

CHAPTER 6

General Discussion



SUMMARY

In my doctoral work, I addressed the question how coding principles of the hippocampal-entorhinal region can support episodic memory and other high-level human cognitive functions. I employed desktop and immersive virtual reality techniques combined with the analysis of multi-voxel patterns to elucidate mechanisms underlying our memory for where and when events take place as well as our ability to imagine future scenarios in front of our mind's eye. Further, based on the combination of a theory from cognitive science and the review of coding principles of the hippocampal-entorhinal region, I developed a theoretical framework for domain-general processing in cognitive spaces. These cognitive spaces offer a representational format for the organization of experience and knowledge in which information is arranged based on its similarity along feature dimensions. The mapping of cognitive spaces via place cell and grid cell population codes provides an explanation for the involvement of the hippocampal-entorhinal region across cognitive domains.

In *Chapter 2*, I employed desktop virtual reality and representational similarity analysis to test the role of spatial codes in imagination. I demonstrated directional coding in human medial temporal lobe structures during mental simulation. Participants imagined directions between buildings in a virtual city and I compared the similarity of multi-voxel patterns between trials as a function of the angular difference between the imagined directions. I observed that pairs of trials in which similar directions were imagined elicited more similar activity patterns in the posterior parahippocampal gyrus. With a resolution of 30°, I demonstrated an absolute directional code at an unprecedented precision with respect to previous fMRI studies investigating representations of heading direction in the human brain (Baumann and Mattingley, 2010; Vass and Epstein, 2013, 2016; Marchette et al., 2014; Chadwick et al., 2015; Shine et al., 2016). Moreover, I showed hexadirectional coding in the posterior-medial entorhinal cortex during imagination. The pattern similarity structure of entorhinal voxels was characterized by a 60° modulation with increased pattern similarity for pairs of trials where the imagined directions were 60° or multiples thereof apart, in line with the notion that grid cell computations support mental simulation. Together, these findings provide evidence for the idea that spatial computations are recruited when draw on our past experience to imagine possible future scenarios. Thereby, my findings suggest a general role for spatial coding principles in the flexible expression of memories.

I further elucidated how putative deviations from canonical spatial coding might affect spatial cognition in a behavioral experiment described in *Chapter 3*. Here, I leveraged immersive virtual reality to demonstrate that environmental geometry deforms human spatial memory in a way that parallels the distortions of rodent grid cell firing patterns in environments with highly polarized geometry (Krupic et al., 2015). The data revealed degraded spatial memory in a trapezoidal compared to a square control environment and, following the distortions of rodent grid cell firing patterns (Krupic et al., 2015) more closely, more pronounced memory degradation in the narrow compared to the broad part of the trapezoid. Importantly, the mnemonic distortions persisted outside of the environment. Following theoretical work on how grid cell computations might underlie the calculation of vectors between locations (Erdem and Hasselmo, 2012; Kubie and Fenton, 2012; Stemmler et al., 2015; Bush et al., 2015; Banino et al., 2018) and how deviations from regular grid patterns might impact these computations (Carpenter and Barry, 2016), I observed differential estimates of identical distances within the different parts of the trapezoid. Within a behavioral study in humans, these findings indirectly suggest the regularity of the grid pattern to be central to its contribution to human spatial memory by highlighting how environmental geometry, shown to distort grid patterns in rodents (Krupic et al., 2015, 2018; Stensola et al., 2015), can distort human spatial cognition.

In *Chapter 4,* I showed that learning a temporal structure shapes the way in which memories are represented in the anterior-lateral subdivision of the entorhinal cortex. Compared to a baseline-scan carried out before the learning task, in which participants acquired knowledge about the positions of objects along a route through a virtual city, multi-voxel patterns changed to represent the temporal relationships between different events in the post-learning scan. From the pattern similarity change from before to after learning, I reconstructed the timeline of events reflecting the encounters of objects along the route. Holistic representations of the learned temporal structure related to memory retrieval in a free recall test. Participants in whom entorhinal pattern similarity change more strongly reflected the temporal structure of the task tended to recall events encountered in temporal proximity at nearby positions. These findings are in line with the report of temporal information carried by populations of neurons in the rodent lateral entorhinal cortex, which might serve as temporal tags for memories (Tsao et al., 2018), and the role of its human homologue region in temporal memory retrieval (Montchal et al., 2019). My findings demonstrate that temporal relationships constitute a key dimension for memory organization in the anterior-lateral entorhinal cortex.

Building upon key spatial coding principles discovered in spatial navigation research and recent evidence for the mapping of non-spatial dimensions by place and grid cells, I proposed in *Chapter 5* that the hippocampal-entorhinal region maps cognitive spaces to support flexible cognition. Cognitive spaces are spanned by feature dimensions satisfying geometric constraints such as betweenness and equidistance (c.f. conceptual spaces in Gärdenfors (2000)). Within these spaces, each stimulus occupies a specific position based on its features. Combining this spatial representational format with growing evidence that place and grid cells also map non-spatial dimensions, I propose that the hippocampal-entorhinal region provides a continuous population code for the dimensions of cognitive spaces. Additional coding principles of this region further support cognitive spaces: The increasing scale of spatial receptive fields along the long-axis of the hippocampal formation (Jung et al., 1994; Hafting et al., 2005; Kjelstrup et al., 2008; Barry et al., 2007; Stensola et al., 2012) suggests multi-scale representations of cognitive spaces at different levels of granularity. The phenomenon of remapping (Muller and Kubie, 1987; Bostock et al., 1991; Leutgeb et al., 2004) highlights the capacity of hippocampal place cells to rapidly map new dimensions and stored maps might be reinstated via attractor dynamics (Wills et al., 2005; Steemers et al., 2016). Conversely, entorhinal grid cells are characterized by rigid firing patterns with stable relationships across different spaces and behavioral states (Fyhn et al., 2007; Yoon et al., 2013; Gardner et al., 2017; Trettel et al., 2017) and might therefore provide a metric for cognitive spaces. Via hippocampal sequential activity, trajectories through cognitive spaces can be simulated and read out for adaptive behavior. Taken together, the theoretical account I outlined combines findings from rodent physiology with concepts from cognitive science to lay out a cognitive neuroscience framework in which spatial coding principles of the hippocampal-entorhinal region constitute domain-general mechanisms for high-level cognition.

In sum, the empirical and theoretical work presented in my doctoral thesis supports the notion that hippocampal-entorhinal coding principles operate across information domains to organize our memory and experience. In the remainder of this final chapter I will discuss the findings from the empirical *Chapters 2-4* in light of the theoretical framework described in *Chapter 5*. Furthermore, I will consider the notion of cognitive spaces in the context of the cognitive map theory (O'Keefe and Nadel, 1978) and the relational memory theory (Eichenbaum and Cohen, 1988, 2014; Eichenbaum et al., 1999); two seminal accounts of hippocampal function. Lastly, I will speculate on how brain areas beyond the hippocampal-entorhinal region may contribute to the processing of cognitive spaces.

HIPPOCAMPAL-ENTORHINAL SEQUENCES SIMULATE TRAJECTORIES THROUGH COGNITIVE SPACES

In *Chapter 5*, I highlighted the importance of hippocampal sequences for the simulation of trajectories through cognitive spaces. Place cell firing sequences corresponding to non-local trajectories have first been observed on linear tracks and subsequently also on more complexes mazes (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Lee and Wilson, 2002; Foster and Wilson, 2006; Diba and Buzsáki, 2007; Davidson et al., 2009; Gupta et al., 2010). Later work has emphasized the sensitivity of sequential hippocampal activity to rewards as well as performance in spatial memory and navigation tasks (Dupret et al., 2010; Jadhav et al., 2012; Singer et al., 2013; Ólafsdóttir et al., 2015, 2017; Ambrose et al., 2016). These findings dovetail with the suggested role of replay in sequence learning in general and the simulation of past trajectories to assign reward value specifically (Foster and Wilson, 2006; Foster and Knierim, 2012; Foster, 2017; Ólafsdóttir et al., 2018). Grid-cell firing in the deep layers of the entorhinal cortex, which receive input from the hippocampus, has been shown to be consistent with trajectories replayed by hippocampal place cells (Ólafsdóttir et al., 2016, 2017), suggesting that not only place cells, but also other spatially tuned cells might be implicated in the simulation of trajectories. Theoretical work has captured these findings to suggest non-local place cell sequences as building blocks for the simulation of future events for decision-making and planning and, more broadly, imagination (Byrne et al., 2007; Buckner, 2010; Bicanski and Burgess, 2018).

This notion dovetails with the central role of medial temporal lobe structures for the human capacity to construct fictitious scenarios (Buckner and Carroll, 2007; Hassabis and Maguire, 2007; Schacter et al., 2007, 2012; Zeidman and Maguire, 2016). For example, patients with hippocampal lesions failed to construct coherent scenes in imagination tasks (Hassabis et al., 2007a) and neuroimaging studies implicate the hippocampus in past and future episodic thinking (Addis et al., 2007; Hassabis et al., 2007b; De Brigard et al., 2013) as well as the recombination of mnemonic details (Barron et al., 2013). Consequently, the hippocampus is thought of as part of a core 'imagination' network of brain regions including medial prefrontal cortex, the precuneus and retrosplenial cortex as well as lateral parietal and lateral temporal cortex (Buckner and Carroll, 2007; Schacter et al., 2007, 2012; Spreng et al., 2009; Benoit and Schacter, 2015). Spatial codes in the hippocampal formation specifically have been put forward to underlie the reinstatement of egocentric views in spatial memory and imagery (Byrne et al., 2007; Bicanski and Burgess, 2018).

The findings reported in *Chapter 2* provide evidence for this notion by demonstrating directional representations in imagination. Absolute directional representations were observed in the parahippocampal gyrus during the imagination period, going beyond similar representations of facing direction in response to visual stimuli (Baumann and Mattingley, 2010; Vass and Epstein, 2013; Chadwick et al., 2015; Shine et al., 2016). Further, the entorhinal cortex exhibited a hexadirectional pattern similarity structure. These representations putatively reflect population activity of head direction and grid cells, respectively. In models of spatial memory and imagery (Byrne et al., 2007; Bicanski and Burgess, 2018), egocentric representations in parietal cortex are transformed via the retrosplenial cortex (which receives head direction input) into allocentric representations in the medial temporal lobe during encoding. Next to place cells, these allocentric representations include boundary and object vector cells (Lever et al., 2009; Høydal et al., 2018). During recall or imagination, the information flow in the network reverses with allocentric representations in the medial temporal lobe driving episodic reconstructions. In this model, head direction information is required for the reconstruction of egocentric parietal representations (Bicanski and Burgess, 2018), in line with the absolute directional representations I demonstrate in the parahippocampal gyrus when participants imagined spatial views between locations. Further, this model of mental imagery ascribes grid cells a key role in navigational planning and mental navigation (Bicanski and Burgess, 2018). Hexadirectional coding putatively reflecting population activity of the entorhinal grid system, which I presented in *Chapter 2* and that is also reported by Horner et al. (2016), might enable shifting the viewpoint of the mind's eye between different locations (Bicanski and Burgess, 2018). In line with this notion, the combination of hexadirectional modulations of theta power during navigation (Chen et al., 2018; Maidenbaum et al., 2018) and increased theta phase coupling between the medial prefrontal cortex and the posterior medial temporal lobe and retrosplenial cortex during dynamic mental imagery (Kaplan et al., 2016) can be interpreted as evidence pointing towards a role of grid-cell computations in the updating of imagined viewpoints.

Building upon these findings and ideas, it is conceivable that the simulation of trajectories through cognitive spaces recruits directional representations as well. As put forward for mental navigation (Bicanski and Burgess, 2018), grid cell firing could support place cell trajectories through cognitive spaces spanned by abstract feature dimensions. For example, computations akin to vector navigation (Bush et al., 2015; Banino et al., 2018) might enable the simulation of a trajectory through a cognitive space to approximate a target position defined by a certain combination

of features. Illustrating this with the car example presented in *Chapter 5*, one might be able to determine the mismatch between the sports car advertised on TV and the family van one is actually looking for if these cars are positioned in a cognitive space spanned by the dimensions of price and available seats, mapped by place and grid cells. This might enable the simulation of the trajectory between them to consider more suitable cars, which cost less and offer more seats. While I reviewed general evidence for simulations of trajectories through cognitive spaces in *Chapter 5*, more specific evidence for spatial coding during simulated trajectories through cognitive spaces remains elusive and poses an intriguing question for future research.

More recently, the content of hippocampal replay has been observed to differ according to task demands, suggesting hippocampal sequences to contribute to planning and decision-making in periods of task engagement and to memory consolidation during rest (Ólafsdóttir et al., 2017, 2018). With respect to the proposed role of replay in memory consolidation (Wilson and McNaughton, 1994; Kumaran and McClelland, 2012; Kumaran et al., 2016b; Ólafsdóttir et al., 2018), it would be interesting to bridge the gap between hippocampal sequences and spontaneous reactivations of learned material that have been linked to enhanced subsequent memory in humans (Deuker et al., 2013; Staresina et al., 2013; Tambini and Davachi, 2013; Schlichting and Preston, 2014; Gruber et al., 2016; Schapiro et al., 2018). To date these studies have focused on isolated reactivations rather than sequential replay, but recent work suggests that the detection of fast sequences might be possible with fMRI (Schuck and Niv, 2018), thereby opening up the possibility of studying the role of sequential reactivations, in line with replay of spatially tuned cells, in human memory consolidation.

This methodological advance might further open up the opportunity to study the involvement of hippocampal sequences in planning using fMRI; following the suggestion in *Chapter 5* that sequential hippocampal activity simulates possible trajectories through cognitive space. For these sequences to support planning and decision-making, the hippocampus might act in concert with striatal and frontal brain areas. Striatal neurons might encode expected rewards based on which hippocampal simulations could be evaluated. Intriguingly, ventral striatal reward signals have been observed during vicarious trial and error at decision points, where hippocampal sequences have been shown to simulate possible trajectories (van der Meer and Redish, 2009; van der Meer et al., 2010; Redish, 2016). This is in line with the proposal that ventral striatal neurons provide covert reward signals for model-based behavior (van der Meer et al., 2010; Redish, 2016).

These reward signals might reflect expected outcomes of trajectories through cognitive space, simulated via sequential hippocampal activity, which might be drawn upon by frontal brain regions to select appropriate actions (Wikenheiser and Schoenbaum, 2016; Sharpe et al., 2019). Wikenheiser and Schoenbaum (2016) suggest the orbitofrontal cortex to play a crucial role in action selection. The selection of appropriate actions requires the representation of the task's structure including its reward contingencies. Consistently, representations of hidden task states have been observed in the orbitofrontal cortex (Wilson et al., 2014; Schuck et al., 2015b, 2016; Sharpe et al., 2019), which, as discussed above, is known to exhibit grid-like coding in humans (Doeller et al., 2010; Jacobs et al., 2013; Constantinescu et al., 2016). Furthermore, prefrontal regions are sensitive to the difficulty of spatial decisions and functional connectivity between medial prefrontal cortex and the hippocampus is enhanced preceding correct choices (Kaplan et al., 2017b).

Taken together, these data suggest prefrontal-hippocampal interactions to support sequential planning (Kaplan et al., 2017a), illustrating how the mapping of experience in cognitive spaces described in *Chapter 5* can be harnessed to guide adaptive future decisions. Consistently, the data presented in *Chapter 2* provide evidence for the involvement of spatial coding mechanisms in mental simulation for navigational planning and the involvement of these representations in planning trajectories through abstract cognitive spaces could be studied using similar methodology.

DEFORMING THE METRIC OF COGNITIVE SPACES

Since their discovery (Hafting et al., 2005), grid cells have been suggested to support path integration through the mapping of navigable space with their regular firing patterns (Hafting et al., 2005; McNaughton et al., 2006; Fiete et al., 2008; Burak and Fiete, 2009; Erdem and Hasselmo, 2012; Mathis et al., 2012; Bush et al., 2015; Moser et al., 2017; Herz et al., 2017). Compared to place cells, grid cell firing is more rigid in that spatial and temporal relations between cells within a module are maintained across environments (Fyhn et al., 2007; Yoon et al., 2013) and behavioral states (Gardner et al., 2017; Trettel et al., 2017). Theoretical work highlights how, despite the periodic firing pattern of each cell, grid firing can encode an agent's position in space and how this spatial representation benefits from the combination of grid patterns across multiple modules (Mathis et al., 2012; Stemmler et al., 2015; Bush et al., 2015; Herz et al., 2017), which differ in the

size of firing fields and the spacing between fields (Barry et al., 2007; Brun et al., 2008; Stensola et al., 2012). Combining these findings and ideas with evidence for grid coding beyond navigable space (Killian et al., 2012; Kraus et al., 2015; Constantinescu et al., 2016; Aronov et al., 2017; Julian et al., 2018; Nau et al., 2018a), I describe in *Chapter 5* that grid cells might function in a similar fashion as a metric for cognitive spaces. Their rigid firing properties might provide a consistent representation of relations between different positions in cognitive space.

Interestingly, deviations from the regular, six-fold symmetric firing patterns have been observed in rodent grid cells (Barry et al., 2007; Krupic et al., 2015, 2018; Stensola et al., 2015), suggesting the grid metric might be distorted under some circumstances. Grid firing patterns align to the walls of the environment (Krupic et al., 2015; Stensola et al., 2015; Sun et al., 2015) and follow compressions and elongations of a familiar rectangular enclosure (Barry et al., 2007). Strikingly, the regularity of the grid firing pattern is distorted in a trapezoidal environment (Krupic et al., 2015) and these distortions appear to be controlled by the geometry of the environment as evidenced by larger distortions near one slanted wall of a quadrilateral environment (Krupic et al., 2018). I investigated the impact of environmental geometry on spatial cognition in a behavioral experiment using immersive virtual reality in *Chapter 3*. My findings demonstrate deformations of spatial cognition, paralleling the distortions of rodent grid cell firing patterns in a trapezoid, which persist in memory outside of the encoding environment. Together with evidence for the sensitivity of human path integration (Chen et al., 2015) and grid-like coding (Navarro Schröder et al., 2017; Nadasdy et al., 2017) to environmental features, this suggests that human spatial cognition might be subject to distortions through environmental geometry, akin to the distortions of rodent grid cell firing patterns.

Connecting my findings with the role of the grid system as a metric for cognitive spaces in *Chapter 5*, the mapping of cognitive spaces might be affected by degraded grid patterns under certain circumstances as well. Correlated feature dimensions or regions of feature combinations that are never observed in a given space might have a similar effect as highly polarized environmental geometry on the way grid cells map these spaces. For example, in the bird feature space spanned by the dimensions of neck and leg length, in which hexadirectional coding in the entorhinal cortex was observed as a function of the orientation of trajectories through the space (Constantinescu et al., 2016), non-orthogonal dimensions would result in deformations of the space. Introducing a correlation between neck and leg length of

the birds could create a triangular space of possible exemplars. Understanding how the entorhinal grid system maps spaces with non-orthogonal dimensions and if grid patterns are subject to similar distortions as in environments with highly polarized geometry under these circumstances will increase our understanding of distance coding in cognitive spaces. Using spatial navigation and memory as a model system, my findings in *Chapter 3* provide first evidence, on the behavioral level, for how distorted grid codes could result in deformations of cognitive processes.

MAPPING TIME FOR MEMORY

While *Chapter 2 and 3* focused on spatial coding, *Chapter 4* addressed temporal mapping in the human entorhinal cortex. Time is a fundamental dimension inherent to our experience and information about when events occurred constitutes a central component of episodic memory (Tulving, 1983, 2002; Wang and Diana, 2017; Ekstrom and Ranganath, 2017; Ranganath, 2018; Howard, 2018). Ever since the famous patient H.M., the hippocampus and surrounding medial temporal lobe structures have been considered crucial for episodic memory (Scoville and Milner, 1957). Consistently, neuroimaging research has corroborated the involvement of the hippocampus and entorhinal cortex in the formation and retrieval of episodic memories (see e.g. Schacter and Wagner (1999); Cabeza and Nyberg (2000); Fernández et al. (1999); Paller and Wagner (2002); Zeidman and Maguire (2016)). In the framework of cognitive spaces, time and space constitute two fundamental dimensions along which experience is organized; similar to proposals putting space and time forward as a scaffold for events to be transformed into episodic memories (Konkel and Cohen, 2009; Schiller et al., 2015; Ekstrom and Ranganath, 2017). In *Chapter 4*, I addressed the question how learning a temporal event structure shapes memories in the entorhinal cortex. The findings demonstrate that the anterior-lateral subregion, considered the human homologue region of the rodent lateral entorhinal cortex (Navarro Schröder et al., 2015; Maass et al., 2015), maps the temporal relationships of events, which have been encountered repeatedly at fixed points along a route.

The region of interest for this analysis was based on the finding that time can be decoded from population activity in the lateral entorhinal cortex of navigating rodents (Tsao et al., 2018). This temporal information is suggested to arise via the encoding of constantly changing experience, which reflects both internal and external states. Constraining experience by introducing a repetitive task in which

the animal ran repeated laps on a maze resulted in decreased decoding accuracy for time across trials, but enhanced temporal decoding within a lap (Tsao et al., 2018). Such a population code in the anterior-lateral entorhinal cortex used to tag memories could explain why the multi-voxel pattern similarity change from before to after learning resembled the temporal structure of the events encountered along the route. As discussed in *Chapter 5*, hippocampal neurons — dubbed time cells — fire consistently at fixed time points during temporal delays and overlap with populations of place and grid cells (MacDonald et al., 2011; Pastalkova et al., 2008; Kraus et al., 2013, 2015). Time cell firing has only been demonstrated for short delays that are repeatedly experienced. One potential mechanism for the generation of temporal codes in the hippocampus could be the input of inherently generated temporal information from the lateral entorhinal cortex (Tsao et al., 2018).

The data I describe in *Chapter 4* suggest holistic representations of temporal relations between events in the anterior-lateral entorhinal cortex. Representational change in the entorhinal cortex reflected the temporal distances between events. The observation of a temporal map in the anterior-lateral entorhinal cortex refines our understanding of the way mnemonic representations are organized in this entorhinal subregion, which has recently been implicated in the retrieval of temporal information (Montchal et al., 2019). One putative interpretation of these findings is that the anterior-lateral entorhinal cortex represents the temporal structure of our experience; in line with the notion put forward in *Chapter 5* that the rigid firing properties of the entorhinal cortex provide a stable metric for cognitive spaces on which more flexible hippocampal codes operate. Similar ideas about structural entorhinal codes, which are drawn upon by conjunctive hippocampal representations have been put forward elsewhere (Behrens et al., 2018; Whittington et al., 2018). Notably, the mapping of temporal structure was selective to the human homologue region of the rodent lateral entorhinal cortex, allowing the speculation that, despite the absence of grid cells in this subregion, the encoding of the structure of experience is a general property of entorhinal representations; potentially going beyond its role in relaying sensory information to the hippocampus (Behrens et al., 2018; Whittington et al., 2018). Previous analyses of the data from *Chapter 4* had revealed that hippocampal pattern similarity change reflected subjectively remembered relations between events in both the temporal and the spatial dimensions, which were dissociated in our experimental design through the use of teleporters (Deuker et al., 2016). Further, participants' memory for relations in the spatial and temporal dimension were influenced by the respective other dimension and hippocampal representations were strongly modulated by the conjunction of remembered spatial

and temporal distances between events (Deuker et al., 2016). Hence, hippocampal representations in this task reflected integrated memories of spatial and temporal relations, which are correlated in everyday experience, whereas the anterior-lateral entorhinal cortex mapped the temporal event structure specifically.

COGNITIVE SPACES IN THE CONTEXT OF COGNITIVE MAPS AND RELATIONAL NETWORKS

The proposal of cognitive spaces in the hippocampal-entorhinal region presented in *Chapter 5* fuses mechanisms discovered in neuroscientific navigation research with a theory from cognitive science. How does it relate to other theories of how past experience guides behavior in general and hippocampal function more specifically?

Based on behavioral experiments in rodents, Tolman (1948) suggested that animals do not only learn stimulus-response associations, but rather form comprehensive representations of their environments. For these he coined the term cognitive maps and he speculated these maps to not only guide navigation behavior, but to also underlie psychological functions. He hypothesized that the failure to learn broad, comprehensive cognitive maps could explain psychological dysfunctions (Tolman, 1948). Tolman's ideas were highly controversial at the time as they constituted a departure from pure behaviorism. The ground-breaking discovery of place cells in the hippocampus of freely moving rats (O'Keefe and Dostrovsky, 1971) provided the first evidence for how cognitive maps of space are encoded in the brain. In their seminal book, O'Keefe and Nadel (1978) proposed an allocentric spatial processing framework — the cognitive map — to reside in the hippocampus, with place cells as its central neural substrate. In an extensive review of the existing literature, O'Keefe and Nadel argued that effects of hippocampal lesions in a wide variety of tasks, including for example discrimination and maze learning, arise from deficits in place learning and exploratory behavior. They extended their ideas from animal models to humans; attributing deficits in spatial mapping and episodic memory in patients with Korsakoff's syndrome or resections of temporal lobe structures to compromised cognitive maps (O'Keefe and Nadel, 1978). Further, they emphasized a functional dissociation between the hemispheres. Whereas the right hippocampus was hypothesized to function similarly to the rodent hippocampus and store episodes in spatio-temporal context, the left hippocampus was suggested to house semantic maps for linguistic processing (O'Keefe and Nadel, 1978, 1979).

The discovery of place cells and the proposal of a cognitive map in the hippocampus sparked decades of research on how the hippocampus represents space for navigation, resulting in the discovery of different types of functionally defined cells and descriptions of the ways in which they encode spatial information (see Moser et al. (2017) for a comprehensive review including a historical perspective). Beyond the encoding of navigable space, the hippocampus has been demonstrated to be sensitive to items or events encountered in the environment (Young et al., 1994; Gothard et al., 1996b; Wood et al., 1999) as well as to exhibit diverging spatial codes for overlapping path segments based on different internal states of the animal (Wood et al., 2000). These findings have been captured in the relational memory theory (Eichenbaum and Cohen, 1988, 2014; Eichenbaum et al., 1999), which proposes a hippocampal memory space in which associative networks of different nodes are formed. In these networks, associative binding creates links between nodes that mirror the relations of items in memory. Experiments conducted in the realm of the relational memory theory have significantly contributed to our understanding how the hippocampus can support associative and transitive inference between overlapping associations. For example, after associating stimuli A and B as well as B and C, relational networks can explain why we can infer that A and C belong together via the shared association of these nodes with node B (Eichenbaum and Cohen, 1988, 2014; Eichenbaum et al., 1999). Consistently, transitive inferences about items in a hierarchy depend on the hippocampus in rodents (Bunsey and Eichenbaum, 1996; Dusek and Eichenbaum, 1997) and the human hippocampus supports in inferential reasoning specifically and the combination of information across episodes in general (Heckers et al., 2004; Preston et al., 2004; Shohamy and Wagner, 2008; Kumaran et al., 2012; Zeithamova et al., 2012a, 2012b; Collin et al., 2015; Milivojevic et al., 2015). Using layer-specific fMRI, recurrence in the hippocampal-entorhinal region has recently been suggested to underlie associative inference. Based on a cue for stimulus A, the association between stimuli A and B is retrieved in the hippocampus and relayed to the deep entorhinal layers. From there it is circulated back to the hippocampus via the superficial entorhinal layers, thereby enabling the retrieval of stimulus C, which had previously been linked to stimulus B, but not A (Koster et al., 2018).

The proposal of cognitive spaces supported by the hippocampal-entorhinal region presented in *Chapter 5* bears similarities and differences in comparison with the cognitive map theory as well as the relational memory theory. The cognitive map theory has a longstanding influence on how we consider space to be represented in the hippocampus and many of the ideas put forward in *Chapter 5* rest upon

40 years of research on how space is mapped in the brain after the original publication by O'Keefe and Nadel (1978) who suggested cognitive maps to reside in the hippocampus. These insights, for example regarding the remapping of place cells, the mapping of space by entorhinal grid cells and the gradient in spatial representation along the hippocampal long-axis, are incorporated in *Chapter 5*; allowing a more detailed formulation of how the spatial processing machinery in the hippocampal formation supports cognition. One important difference compared to the cognitive map theory and the relational memory theory, which focus on the hippocampus, is that grid cells in the entorhinal cortex are ascribed a central role for cognitive spaces. Going beyond the level of analogy for cognitive maps in memory and linguistic processing, the account outlined in *Chapter 5* specifies how place and grid cells provide a continuous population code for the geometrically constrained dimensions spanning cognitive spaces. While multi-dimensional spaces for experience have also been put forward in the context of the relational memory theory (Konkel and Cohen, 2009), this account has focused on associative networks of different items within these spaces rather than a formal description of the space itself. The geometric constraints on the dimensions of cognitive spaces are key to the framework described in *Chapter 5*.

The dimensions of cognitive spaces are constrained to satisfy the geometric properties of betweenness and equidistance. Betweenness implies that for a set of points (A, B, C), where B lies in between A and C, C cannot be between B and A. Further, betweenness enables inference: If B is between A and C and C lies between B and D, then C must be between A and D. A second geometric constraint imposed on the dimensions of cognitive spaces is equidistance. This permits statements about the relations between points in cognitive space. If the distance between A and B is equal to the distance between C and D and we learn that E and F are the same distance from each other as A and B, then E and F must also be separated by the same distance as C and D. Combining the constraints of betweenness and equidistance further allows propositions such as that the distance between A and C must be the sum of the distances between points A and B and points B and C, if B is in between A and C. A third assumption about the dimensions is that they have an underlying distance function; the Euclidean or the city block metric for example. By ascribing a metric or distance function to the axes, similarity between arbitrary positions naturally arises as a function of the distance between these positions in this space (Gärdenfors, 2000). At the heart of the account put forward in *Chapter 5* is the idea that the continuous population code of place and grid cells maps the dimensions of these geometrically defined cognitive spaces. This idea reconciles a

growing body of evidence demonstrating that the hippocampal-entorhinal region in general and place and grid cells in particular map dimensions of experience (Killian et al., 2012; Kraus et al., 2013, 2015; Constantinescu et al., 2016; Aronov et al., 2017; Danjo et al., 2018; Omer et al., 2018). Hence, at the level of neural circuits, similarity between stimuli can be captured via place and grid cell population activity encoding these positions.

Contrarily, similarity between stimuli or events is conceptualized as links between nodes in relational networks (Eichenbaum and Cohen, 1988, 2014; Eichenbaum et al., 1999). This constitutes an important difference in how similarity between stimuli arises, which extends to how cognitive and behavioral processes can support generalization to novel stimuli (Shepard, 1987). The notion of properties and concepts as convex regions (Gärdenfors, 2000) in cognitive space, which rests upon the geometric constraints on the dimensions, can be harnessed for generalization to novel situations and never-experienced stimuli. For example, if memory contains two events, one where an action A1 led to a result R1 that did not quite reach a goal and one where an action A2 had a result R2 that overshot the goal, then by convexity an action A3 between A1 and A2 can be found where the result R3 will be between R1 and R2 and thus maybe come closer to the goal. In the relational memory theory, inference is limited to nodes that share associations and therefore relational networks cannot easily explain inference to completely novel stimuli that have not been explicitly associated with any network node.

Taken together, the ideas put forward in *Chapter 5* build upon the seminal empirical findings and theoretical advances made in the context of the cognitive map theory as well as the relational memory theory. However, they incorporate exciting recent discoveries about how place and grid cells in the hippocampal-entorhinal region map dimensions of experience beyond space for navigation to provide a detailed account of a spatial representational format for flexible cognition.

OUTLOOK: COGNITIVE SPACES BEYOND THE HIPPOCAMPAL-ENTORHINAL REGION?

The proposal for cognitive spaces in *Chapter 5* focuses on processing mechanisms in the hippocampal-entorhinal region. However, the hippocampus is thought to operate as part of a core network of brain regions that contributes to a wider variety of cognitive functions. In the context of navigation, this network has been

described to consist of medial temporal lobe regions including the hippocampus and the parahippocampal cortex, the retrosplenial complex (RSC) consisting of retrosplenial and posterior cingulate cortices and the posterior parietal cortex (Spiers and Maguire, 2006; Epstein et al., 2017). Within this network, the hippocampus has classically been viewed as holding allocentric spatial representations, whereas the posterior parietal cortex has been assumed to operate in a more egocentric reference frame (Byrne et al., 2007; Ekstrom et al., 2017). The RSC has been suggested to translate between these reference frames (Byrne et al., 2007; Vann et al., 2009; Clark et al., 2018), in line with directional coding in local reference frames observed in both humans (Marchette et al., 2014) and rodents (Jacob et al., 2017). In *Chapter 5* specifically and throughout this thesis in general, I have argued for domain-general codes in the hippocampal-entorhinal region. However, other brain areas of the navigation network outlined above have also been suggested to operate independently of specific sensory modalities (Wolbers et al., 2011; Vass and Epstein, 2016; Ekstrom et al., 2017; Auger and Maguire, 2018a). Below I speculate how processing in the retrosplenial complex and the parahippocampal cortex could contribute to the encoding of particularly relevant stimuli. The spatial coding principles described in *Chapter 5* are suggested to operate across cognitive domains, but importantly they should not be seen as coding principles governing the entire brain. To illustrate this, I will summarize a study on face processing that demonstrates different neural codes for a multi-dimensional stimulus space in the inferotemporal cortex in monkeys (Chang and Tsao, 2017). However, there is evidence for grid-like coding beyond the entorhinal cortex in humans (Doeller et al., 2010; Jacobs et al., 2013; Constantinescu et al., 2016). I will briefly discuss spatial codes beyond the hippocampal-entorhinal region and a recent proposal for how grid coding might reflect prior knowledge. I will also touch upon frontal-hippocampal interactions in the context of schemas, which are thought of as a prior knowledge structures influencing information processing from perception to memory retrieval.

Speculative roles for the parahippocampal place area and the retrosplenial complex in cognitive space

The parahippocampal place area (PPA) is a functionally defined brain region extending from the posterior parahippocampal cortex to the lingual and fusiform gyri as well as the collateral sulcus (Epstein and Kanwisher, 1998; Epstein, 2014). It is thought to be involved in the processing of scene geometry and the finding of stronger scene-related activity in contrast to object-related activity in brain areas such as the PPA constitutes a reliable finding in human fMRI (Julian et al., 2012;

Epstein, 2014). The notion that PPA encodes the spatial layout (or geometry) of scenes enjoys wide support in the field (summarized e.g. in Epstein (2014)). For example, the PPA has been shown to be sensitive to the spatial expanse as well as the arrangement of boundaries in visual scenes, prominent features of scene geometry (Kravitz et al., 2011; Park et al., 2011). However the precise nature of its neural coding mechanisms remain elusive (Epstein and Kanwisher, 1998; Epstein, 2014). While PPA activation in response to haptic exploration of scenes (Wolbers et al., 2011) suggests at least some degree of modality-independent processing, this brain area has mostly been studied in the visual domain. Beyond scene geometry, studies have revealed that the PPA preferentially encodes navigationally relevant stimuli (Janzen and van Turennout, 2004) and that the identity of landmarks is encoded in a way that generalizes across different, visually distinct images of a given landmark (Marchette et al., 2015). Hence, one might speculate the PPA to carry information about stimuli at particularly relevant positions in cognitive space based on its increased activation in response to objects at decision points (Janzen and van Turennout, 2004).

In the context of navigation, the RSC has often been ascribed a role in the translation between allocentric representations in the medial temporal lobe and egocentric representations in the posterior parietal cortex (Byrne et al., 2007; Vann et al., 2009; Bicanski and Burgess, 2018; Clark et al., 2018). Neurons in the RSC encode the position of an animal along a route, the direction of turns and heading directions in different reference frames (Alexander and Nitz, 2015, 2017; Jacob et al., 2017; Marchette et al., 2014). Nau et al. (2018b) have recently proposed the RSC to play a similar role in the transformation from retinotopic coordinates in the parietal cortex to world-centered coordinates in the hippocampal-entorhinal region for the mapping of visual space. Likewise, world-centered representations of visual space, particularly grid-like codes (Killian et al., 2012; Julian et al., 2018; Nau et al., 2018a), might play a role in transsaccadic memory as a way of anchoring visual receptive fields to the external world also in cortical regions (Nau et al., 2018b). Based on recent work on the sensitivity of the RSC to the stability of landmarks (Auger et al., 2015; Auger and Maguire, 2018b), one might further speculate that the RSC might be involved in mapping positions and trajectories relative to particularly relevant stimuli in cognitive space, potentially in line with its role in reference frame transformations during navigation (Vann et al., 2009; Epstein et al., 2017; Clark et al., 2018). Notably, the RSC is sensitive to stability beyond the spatial domain as indexed by increased activity for verbal descriptions of permanent, habitual compared to transient actions (Auger and Maguire, 2018a); providing some support for the

speculation that it might encode reference positions in cognitive space.

Face space: A counter-example

Importantly, the spatial coding principles outlined in *Chapter 5* should not be seen as a blueprint for processing that can simply be transferred to regions throughout the entire brain. This is illustrated for example by a recent study investigating the coding properties of neurons in face-responsive areas of monkey inferotemporal cortex (Chang and Tsao, 2017). The authors characterized neural responses to face stimuli varying along 50 dimensions. Individual neurons did not encode positions in this high-dimensional face space, but selectively responded to manipulations of a preferred dimension of the face space. Specifically, the neurons exhibited ramp-like tuning characterized by minimum and maximum firing at extreme feature values of the preferred dimension, while not being influenced by manipulations of non-preferred dimensions (Chang and Tsao, 2017). This constitutes a marked difference from the spatially constrained firing of place and grid cells and suggests that while brain areas central to face processing might also map stimuli into high-dimensional spaces, the underlying neural code might represent these spaces in a very different way. A possible additional difference might be the domain-generality of constrained encoding of relevant dimensions in the hippocampal formation, though neither hippocampal activity nor neural responses in inferotemporal cortex to non-face visual stimuli were recorded in the aforementioned study.

Spatial codes for prior knowledge?

While the description of cognitive spaces in *Chapter 5* focused on place and grid cells in the hippocampal-entorhinal region, evidence for the existence of these cell types has been reported in other brain areas. In terms of spatially constrained firing, cells in prefrontal cortex exhibited place fields and these firing fields clustered around goal locations (Hok et al., 2005), in line with the suggested role of prefrontal cortex in planning and decision-making. A similar sensitivity of place cells to goal locations has also been observed in the hippocampus (Hollup et al., 2001; Hok et al., 2007; Dupret et al., 2010). In human fMRI, grid-like hexadirectional signals have been observed in frontal and parietal areas of the human brain during virtual navigation (Doeller et al., 2010) and during trajectories through a cognitive space defined by two bird feature dimensions (Constantinescu et al., 2016). Likewise, direct recordings in presurgical patients have demonstrated the existence of place- and grid-cell-like coding also in the cingulate cortex of the human brain (Jacobs et al., 2013; Qasim et al., 2018). If place and grid cells exist beyond the hippocampal formation, this begs the questions how spatial codes in these regions interact.

Recently, grid cells in frontal brain areas and entorhinal cortex have been argued to reflect knowledge about the structure of two-dimensional spaces, which can be applied to different tasks of the same structure (Behrens et al., 2018). Not unlike the flexible mapping of different dimensions described in *Chapter 5*, spatial navigation is just one instantiation of a task with a continuous two-dimensional structure of possible states in this framework. When taking a reinforcement learning view on the problem of navigation, different positions in space define possible states of the task (Gustafson and Daw, 2011; Behrens et al., 2018) and grid-like basis functions efficiently capture the structure of possible transitions between states of such a two-dimensional space (Dordek et al., 2016; Stachenfeld et al., 2017). This structural information can enable generalizations across different tasks with the same underlying structure, independent of specific sensory material; leading to the intriguing idea that grid cells underlie the central role of ventral prefrontal areas in representing abstract relationships between task events (Behrens et al., 2018; Whittington et al., 2018). With respect to hexadirectional signals in fMRI, Constantinescu et al. (2016) provide some fMRI evidence that entorhinal and frontal grid cells might be operating in concert to support the mapping of abstract task spaces. This goes back to the finding that the orientation of hexadirectional signals was similar between ventromedial prefrontal and entorhinal cortex, suggesting shared coding (Constantinescu et al., 2016). While these findings and ideas are in line with the notion that spatial coding mechanisms in other brain regions might operate in a similar way as in the hippocampal-entorhinal region, the precise interactions between grid coding in frontal and entorhinal cortex are not well understood and spatial codes beyond the hippocampal formation are — in comparison — understudied in general. Here, spatial navigation can continue to serve as an important model system to examine properties of spatial codes in general and grid codes specifically across brain regions to understand in more detail how these codes are generated and transferred between regions and how they interact. This will in turn enable us to elucidate how knowledge about task structures is represented and transferred between regions.

The influence of prior knowledge on the processing, encoding, interpretation, consolidation and retrieval of information has been extensively studied in the realm of schemas (van Kesteren et al., 2012; Gilboa and Marlatte, 2017; Mack et al., 2017; Fernández and Morris, 2018). The formation, instantiation and reinstatement of these abstract knowledge structures is thought to be supported by ventromedial prefrontal regions, the hippocampus and the posterior neocortex including the temporoparietal junction and the angular gyrus (van Kesteren et al., 2012; Preston

and Eichenbaum, 2013; Gilboa and Marlatte, 2017). How precisely the prefrontal cortex and the hippocampus interact remains controversial (van Kesteren et al., 2012; Preston and Eichenbaum, 2013; Gilboa and Marlatte, 2017). On the one hand, the role of the hippocampus has been suggested to lie in the relation of new events to prior knowledge via feature overlap, while distinguishing different contexts (Preston and Eichenbaum, 2013; Mack et al., 2017); akin to the mapping of experience in cognitive spaces based on stimulus features. The prefrontal cortex here stores contextual rules and biases hippocampal processing accordingly (Preston and Eichenbaum, 2013), putatively in line with the retrieval of stored maps for the appropriate cognitive space. Conversely, the ventromedial prefrontal cortex has been suggested to inhibit the hippocampus when incoming information is congruent with an existing schema (van Kesteren et al., 2010, 2012). Here, medial prefrontal cortex is central to rapid memory formation if to-be-learnt information matches an existing schema, while the medial temporal lobe is especially important in the context of novel material that is inconsistent with prior knowledge (van Kesteren et al., 2012). The findings that grid-like coding principles are shared between the hippocampal-entorhinal region and frontal brain areas (Doeller et al., 2010; Jacobs et al., 2013; Constantinescu et al., 2016) emphasizes the importance of understanding the interaction of these brain regions. Studying these interactions in the context of spatial navigation and memory might generate testable hypotheses for the study of these codes in complex schemas, including large knowledge structures (van Kesteren et al., 2014). Further, the temporoparietal junction and angular gyrus specifically are thought to be crucial for the storage of abstract rules (Wagner et al., 2015; Gilboa and Marlatte, 2017; Fernández and Morris, 2018) and the temporoparietal junction exhibits hexadirectional signals during trajectories through cognitive space (Constantinescu et al., 2016). However, one important caveat is that the coding principles described for cognitive spaces in *Chapter 5* have typically been investigated in situations where the underlying stimulus dimensions are highly constrained, whereas high-level schemas are thought to encompass relationships between much more complex experiences, from which intricate knowledge structures are derived. However, the possibility that grid-like coding principles might be to some extent shared across regions important for the representation of abstract prior knowledge is intriguing and might open up a fruitful avenue to study the coding principles underlying highly complex schemas formed from the commonalities of experience.

OUTSTANDING QUESTIONS

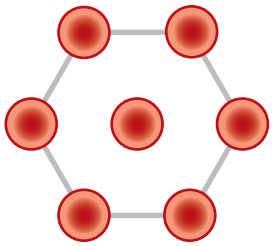
In this last section, I will highlight a few of the outstanding questions that arise from the work presented in this thesis. Regarding the role of spatial representations in imagination demonstrated in *Chapter 2*, it would be of interest to investigate the role of directional codes in mental simulation more generally. For example, can they be observed during tasks beyond navigable space? With respect to the role of mental simulation in planning, the intriguing question remains which trajectories are selected for simulation and how these simulations are evaluated. Turning towards the distortions of spatial cognition observed in *Chapter 3*, the neural mechanisms underlying these distortions have not been studied in humans. For example, these distortions might result in degraded hexadirectional signals in environments with highly polarized environmental geometry or in an altered pattern similarity structure when comparing representations of learned locations in a framework as in *Chapter 4* and Deuker et al. (2016). Capitalizing on the advances in mobile MEG setups might enable the recording of neural signals in combination with immersive VR to investigate distortions in the recently described electrophysiological signatures of hexadirectional coding in the theta frequency band (Chen et al., 2018; Maidenbaum et al., 2018) as well as in broadband high-frequency activity (Staudigl et al., 2018). Understanding the distortions of grid coding will enhance our comprehension of how the entorhinal cortex provides a distance function for cognitive spaces. In terms of the temporal structure of memories represented in the entorhinal cortex that I showed in *Chapter 4*, it would be of interest to further dissociate temporal intervals and sequence order, which are typically highly correlated to understand their unique contributions to temporal representations. A further question concerns the representation of nested temporal scales in the entorhinal cortex and hippocampus, which could be addressed for example by comparing similar temporal intervals across multiple days.

The proposal of cognitive spaces in the hippocampal-entorhinal region opens a large number of questions, which have been touched upon in *Chapter 5* as well as in this general discussion. Of particular interest would be to understand the representation of complex stimuli in cognitive spaces. So far, investigations of abstract task spaces have mostly been restricted to one or two dimensions, whereas our experience is characterized by high-dimensional stimuli. A different question concerns the precise metric underlying cognitive spaces. While a Euclidean metric might seem most intuitive, other distance functions are similarly plausible and some reports suggest topological representations to support navigation (Dabaghian et al.,

2014; Javadi et al., 2017; Warren et al., 2017). Furthermore, the extent to which spatial coding principles are domain-general and function similarly in cognitive space remains to be explored. For example, this includes the recent proposal that the processing of boundaries in spatial navigation might share neural substrates with the detection of event boundaries in episodic memory (Brunec et al., 2018b). Likewise, as exemplified by the speculations on retrosplenial and parahippocampal cortices above, the contribution brain areas acting in concert with the hippocampal-entorhinal region make to cognitive spaces remains to be elucidated in future research.

CONCLUDING REMARKS

In this thesis, I presented experimental and theoretical work addressing the question how coding principles of the hippocampal-entorhinal region support our memory and map experience for flexible cognition. Using virtual reality technology, behavioral experiments and representational similarity analysis of fMRI multi-voxel patterns, I tested hypotheses derived from findings in rodent electrophysiology to elucidate the mechanisms underlying human imagination as well as spatial and temporal memory. Bridging insights into coding mechanisms of the hippocampal-entorhinal region from animal and human research and combining them with a theory from cognitive science, I proposed a framework in which cognitive spaces provide a domain-general format for flexible cognition.



REFERENCES

Addis, D.R., Wong, A.T., and Schacter, D.L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. Neuropsychologia *45*, 1363–1377.

Alexander, A.S., and Nitz, D.A. (2015). Retrosplenial cortex maps the conjunction of internal and external spaces. Nat. Neurosci. *18*, 1143–1151.

Alexander, A.S., and Nitz, D.A. (2017). Spatially Periodic Activation Patterns of Retrosplenial Cortex Encode Route Sub-spaces and Distance Traveled. Curr. Biol. *27*, 1551-1560.e4.

Ambrose, R.E., Pfeiffer, B.E., and Foster, D.J. (2016). Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. Neuron *91*, 1124–1136.

Anderson, J., Jenkinson, M., and Smith, S. (2010). Non-linear registration, aka spatial normalisation. FMRIB Tech. Rep. *TR07JA2*.

Aronov, D., Nevers, R., and Tank, D.W. (2017). Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. Nature *543*, 719–722.

Auger, S.D., and Maguire, E.A. (2018a). Retrosplenial Cortex Indexes Stability beyond the Spatial Domain. J. Neurosci. *38*, 1472–1481.

Auger, S.D., and Maguire, E.A. (2018b). Dissociating Landmark Stability from Orienting Value Using Functional Magnetic Resonance Imaging. J. Cogn. Neurosci. *30*, 698–713.

Auger, S.D., Zeidman, P., and Maguire, E.A. (2015). A central role for the retrosplenial cortex in de novo environmental learning. ELife *4*, e09031.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., and Norman, K.A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. Neuron *95*, 709-721.e5.

Balkenius, C., and Gärdenfors, P. (2016). Spaces in the Brain: From Neurons to Meanings. Front. Psychol. *7*.

Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., Pritzel, A., Chadwick, M.J., Degris, T., Modayil, J., et al. (2018). Vector-based navigation using grid-like representations in artificial agents. Nature *557*, 429–433.

Barnett, A.J., O'Neil, E.B., Watson, H.C., and Lee, A.C.H. (2014). The human hippocampus is sensitive to the durations of events and intervals within a sequence. Neuropsychologia *64*, 1–12.

Barron, H.C., Dolan, R.J., and Behrens, T.E.J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci. *16*, 1492–1498.

Barry, C., Hayman, R., Burgess, N., and Jeffery, K.J. (2007). Experience-dependent rescaling of entorhinal grids. Nat. Neurosci. *10*, 682–684.

Barry, C., Ginzberg, L.L., O'Keefe, J., and Burgess, N. (2012). Grid cell firing patterns signal environmental novelty by expansion. Proc. Natl. Acad. Sci. *109*, 17687–17692.

Baumann, O., and Mattingley, J. (2010). Medial parietal cortex encodes perceived heading direction in humans. J. Neurosci. *30*, 12897–12901.

Behrens, T.E.J., Muller, T.H., Whittington, J.C.R., Mark, S., Baram, A.B., Stachenfeld, K.L., and Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron *100*, 490–509.

Bellmund, J.L.S., Deuker, L., Navarro Schröder, T., and Doeller, C.F. (2016). Grid-cell representations in mental simulation. ELife *5*, e17089.

Bellmund, J.L.S., Gärdenfors, P., Moser, E.I., and Doeller, C.F. (2018a). Navigating cognition: Spatial codes for human thinking. Science *362*, eaat6766.

Bellmund, J.L.S., Deuker, L., and Doeller, C.F. (2018b). Donderstown. Open Sci. Framew.

Benoit, R.G., and Schacter, D.L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. Neuropsychologia *75*, 450–457.
Ben-Yakov, A., Eshel, N., and Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. J. Exp. Psychol. Gen. *142*, 1255–1263.
Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. J. Stat. Softw. *31*.
Bicanski, A., and Burgess, N. (2018). A neural-level model of spatial memory and imagery. ELife *7*.
Bird, C.M., Bisby, J.A., and Burgess, N. (2012). The hippocampus and spatial constraints on mental imagery. Front. Hum. Neurosci. *6*, 142.
Bonnici, H.M., Kumaran, D., Chadwick, M.J., Weiskopf, N., Hassabis, D., and Maguire, E.A. (2012). Decoding representations of scenes in the medial temporal lobes. Hippocampus *22*, 1143–1153.
Bornstein, A.M., and Norman, K.A. (2017). Reinstated episodic context guides sampling-based decisions for reward. Nat. Neurosci. *20*, 997–1003.
Bostock, E., Muller, R.U., and Kubie, J.L. (1991). Experience-dependent modifications of hippocampal place cell firing. Hippocampus *1*, 193–205.
Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., Muñoz, L.D., Mullinger, K.J., Tierney, T.M., Bestmann, S., et al. (2018). Moving magnetoencephalography towards real-world applications with a wearable system. Nature *555*, 657–661.
Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. *10*, 433–436.
Brewer, J.B., Zhao, Z., Desmond, J.E., Glover, G.H., and Gabrieli, J.D.E. (1998). Making Memories: Brain Activity that Predicts How Well Visual Experience Will Be Remembered. Science *281*, 1185–1187.
Brown, T.I., Carr, V.A., LaRocque, K.F., Favila, S.E., Gordon, A.M., Bowles, B., Bailenson, J.N., and Wagner, A.D. (2016). Prospective representation of navigational goals in the human hippocampus. Science *352*, 1323–1326.
Brun, V.H., Solstad, T., Kjelstrup, K.B., Fyhn, M., Witter, M.P., Moser, E.I., and Moser, M.-B. (2008). Progressive increase in grid scale from dorsal to ventral medial entorhinal cortex. Hippocampus *18*, 1200–1212.
Brunec, I.K., Javadi, A.-H., Zisch, F.E.L., and Spiers, H.J. (2017). Contracted time and expanded space: The impact of circumnavigation on judgements of space and time. Cognition *166*, 425–432.
Brunec, I.K., Bellana, B., Ozubko, J.D., Man, V., Robin, J., Liu, Z.-X., Grady, C., Rosenbaum, R.S., Winocur, G., Barense, M.D., et al. (2018a). Multiple Scales of Representation along the Hippocampal Anteroposterior Axis in Humans. Curr. Biol. *28*, 2129-2135.e6.
Brunec, I.K., Moscovitch, M., and Barense, M.D. (2018b). Boundaries Shape Cognitive Representations of Spaces and Events. Trends Cogn. Sci. *22*, 637–650.
Buckner, R.L. (2010). The role of the hippocampus in prediction and imagination. Annu. Rev. Psychol. *61*, 27–48.
Buckner, R.L., and Carroll, D.C. (2007). Self-projection and the brain. Trends Cogn. Sci. *11*, 49–57.
Bullens, J., Nardini, M., Doeller, C.F., Braddick, O., Postma, A., and Burgess, N. (2010). The role of landmarks and boundaries in the development of spatial memory. Dev. Sci. *13*, 170–180.
Bunsey, M., and Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. Nature *379*, 255–257.

Burak, Y., and Fiete, I.R. (2009). Accurate Path Integration in Continuous Attractor Network Models of Grid Cells. PLOS Comput. Biol. *5*, e1000291.
Bush, D., Barry, C., Manson, D., and Burgess, N. (2015). Using Grid Cells for Navigation. Neuron *87*, 507–520.
Buzsáki, G., and Moser, E.I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. Nat. Neurosci. *16*, 130–138.
Byrne, P., Becker, S., and Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psychol. Rev. *114*, 340–375.
Cabeza, R., and Nyberg, L. (2000). Neural bases of learning and memory: functional neuroimaging evidence. Curr. Opin. Neurol. *13*, 415.
Cakmak, T., and Hager, H. (2014). Cyberith virtualizer: a locomotion device for virtual reality. In ACM SIGGRAPH 2014 Emerging Technologies, (ACM), p. 6.
Campbell, M.G., Ocko, S.A., Mallory, C.S., Low, I.I.C., Ganguli, S., and Giocomo, L.M. (2018). Principles governing the integration of landmark and self-motion cues in entorhinal cortical codes for navigation. Nat. Neurosci. *21*, 1096–1106.
Canto, C.B., Wouterlood, F.G., and Witter, M.P. (2008). What Does the Anatomical Organization of the Entorhinal Cortex Tell Us?
Carpenter, F., and Barry, C. (2016). Distorted Grids as a Spatial Label and Metric. Trends Cogn. Sci. *20*, 164–167.
Carpenter, F., Manson, D., Jeffery, K., Burgess, N., and Barry, C. (2015). Grid Cells Form a Global Representation of Connected Environments. Curr. Biol. *25*, 1176–1182.
Carr, M.F., Jadhav, S.P., and Frank, L.M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. Nat. Neurosci. *14*, 147–153.
Chadwick, M.J., Jolly, A.E.J., Amos, D.P., Hassabis, D., and Spiers, H.J. (2015). A Goal Direction Signal in the Human Entorhinal/Subicular Region. Curr. Biol. *25*, 87–92.
Chang, L., and Tsao, D.Y. (2017). The Code for Facial Identity in the Primate Brain. Cell *169*, 1013-1028.e14.
Chen, D., Kunz, L., Wang, W., Zhang, H., Wang, W.-X., Schulze-Bonhage, A., Reinacher, P.C., Zhou, W., Liang, S., Axmacher, N., et al. (2018). Hexadirectional Modulation of Theta Power in Human Entorhinal Cortex during Spatial Navigation. Curr. Biol. *28*, 3310-3315.e4.
Chen, G., King, J.A., Burgess, N., and O'Keefe, J. (2013). How vision and movement combine in the hippocampal place code. Proc. Natl. Acad. Sci. *110*, 378–383.
Chen, X., He, Q., Kelly, J.W., Fiete, I.R., and McNamara, T.P. (2015). Bias in Human Path Integration Is Predicted by Properties of Grid Cells. Curr. Biol. *25*, 1771–1776.
Clark, R.E., and Squire, L.R. (2013). Similarity in form and function of the hippocampus in rodents, monkeys, and humans. Proc. Natl. Acad. Sci. *110*, 10365–10370.
Clark, B.J., Simmons, C.M., Berkowitz, L.E., and Wilber, A.A. (2018). The retrosplenial-parietal network and reference frame coordination for spatial navigation. Behav. Neurosci.
Cohen, N.J., and Eichenbaum, H. (1993). Memory, amnesia, and the hippocampal system. (Cambridge, MA, US: The MIT Press).
Cohen, N.J., and Squire, L.R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. Science *210*, 207–210.
Collin, S.H.P., Milivojevic, B., and Doeller, C.F. (2015). Memory hierarchies map onto the hippocampal long axis in humans. Nat. Neurosci. *18*, 1562–1564.
Collins, A.M., and Quillian, M.R. (1969). Retrieval time from semantic memory. J. Verbal Learn. Verbal Behav. *8*, 240–247.
Constantinescu, A.O., O'Reilly, J.X., and Behrens, T.E.J. (2016). Organizing conceptual knowledge in humans with a gridlike code. Science *352*, 1464–1468.

Copara, M.S., Hassan, A.S., Kyle, C.T., Libby, L.A., Ranganath, C., and Ekstrom, A.D. (2014). Complementary Roles of Human Hippocampal Subregions during Retrieval of Spatiotemporal Context. J. Neurosci. *34*, 6834–6842.
Cullen, K.E., and Taube, J.S. (2017). Our sense of direction: progress, controversies and challenges. Nat. Neurosci. *20*, 1465.
Dabaghian, Y., Brandt, V.L., and Frank, L.M. (2014). Reconceiving the hippocampal map as a topological template. ELife *3*, e03476.
Danjo, T., Toyoizumi, T., and Fujisawa, S. (2018). Spatial representations of self and other in the hippocampus. Science *359*, 213–218.
Davachi, L., and DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. Trends Cogn. Sci. *19*, 92–99.
Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Hippocampal Replay of Extended Experience. Neuron *63*, 497–507.
De Almeida, L., Idiart, M., Villavicencio, A., and Lisman, J. (2012). Alternating Predictive and Short-term memory modes of Entorhinal Grid cells. Hippocampus *22*, 1647–1651.
De Brigard, F., Addis, D.R., Ford, J.H., Schacter, D.L., and Giovanello, K.S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. Neuropsychologia *51*, 2401–2414.
Deshmukh, S.S., and Knierim, J.J. (2011). Representation of Non-Spatial and Spatial Information in the Lateral Entorhinal Cortex. Front. Behav. Neurosci. *5*.
Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage *31*, 968–980.
Deuker, L., Olligs, J., Fell, J., Kranz, T.A., Mormann, F., Montag, C., Reuter, M., Elger, C.E., and Axmacher, N. (2013). Memory Consolidation by Replay of Stimulus-Specific Neural Activity. J. Neurosci. *33*, 19373–19383.
Deuker, L., Bellmund, J.L.S., Navarro Schröder, T., and Doeller, C.F. (2016). An event map of memory space in the hippocampus. ELife *5*, e16534.
Diba, K., and Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. Nat. Neurosci. *10*, 1241–1242.
Doeller, C.F., and Burgess, N. (2008). Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. Proc. Natl. Acad. Sci. U. S. A. *105*, 5909–5914.
Doeller, C.F., King, J. a, and Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. Proc. Natl. Acad. Sci. U. S. A. *105*, 5915–5920.
Doeller, C.F., Barry, C., and Burgess, N. (2010). Evidence for grid cells in a human memory network. Nature *463*, 657–661.
Doll, B.B., Duncan, K.D., Simon, D.A., Shohamy, D., and Daw, N.D. (2015). Model-based choices involve prospective neural activity. Nat. Neurosci. *18*, 767–772.
Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron *56*, 43–57.
Donato, F., Jacobsen, R.I., Moser, M.-B., and Moser, E.I. (2017). Stellate cells drive maturation of the entorhinal-hippocampal circuit. Science *355*, eaai8178.
Dordek, Y., Soudry, D., Meir, R., and Derdikman, D. (2016). Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. ELife *5*, e10094.

Dragoi, G., and Tonegawa, S. (2011). Preplay of future place cell sequences by hippocampal cellular assemblies. Nature *469*, 397–401.
DuBrow, S., and Davachi, L. (2014). Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. J. Neurosci. *34*, 13998–14005.
DuBrow, S., and Davachi, L. (2016). Temporal binding within and across events. Neurobiol. Learn. Mem. *134*, 107–114.
Dupret, D., O'Neill, J., Pleydell-Bouverie, B., and Csicsvari, J. (2010). The reorganization and reactivation of hippocampal maps predict spatial memory performance. Nat. Neurosci. *13*, 995–1002.
Dusek, J.A., and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. Proc. Natl. Acad. Sci. *94*, 7109–7114.
Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. Nat. Rev. Neurosci. *15*, 732–744.
Eichenbaum, H., and Cohen, N.J. (1988). Representation in the hippocampus: what do hippocampal neurons code? Trends Neurosci. *11*, 244–248.
Eichenbaum, H., and Cohen, N.J. (2014). Can We Reconcile the Declarative Memory and Spatial Navigation Views on Hippocampal Function? Neuron *83*, 764–770.
Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? Neuron *23*, 209–226.
Eickhoff, S.B., Paus, T., Caspers, S., Grosbras, M.-H., Evans, A.C., Zilles, K., and Amunts, K. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. NeuroImage *36*, 511–521.
Ekstrom, A.D., and Ranganath, C. (2017). Space, time, and episodic memory: The hippocampus is all over the cognitive map. Hippocampus.
Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T.A., Isham, E.A., Newman, E.L., and Fried, I. (2003). Cellular networks underlying human spatial navigation. Nature *425*, 184–188.
Ekstrom, A.D., Huffman, D.J., and Starrett, M. (2017). Interacting networks of brain regions underlie human spatial navigation: a review and novel synthesis of the literature. J. Neurophysiol. *118*, 3328–3344.
Epstein, R.A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. *12*, 388–396.
Epstein, R.A. (2014). Neural systems for visual scene recognition. In Scene Vision: Making Sense of What We See, K. Kveraga, and Bar, eds. (Cambridge, MA, US: MIT Press), pp. 105–134.
Epstein, R., and Kanwisher, N. (1998). A cortical representation of the local visual environment. Nature *392*, 598–601.
Epstein, R.A., Patai, E.Z., Julian, J.B., and Spiers, H.J. (2017). The cognitive map in humans: spatial navigation and beyond. Nat. Neurosci. *20*, 1504.
Erdem, U.M., and Hasselmo, M. (2012). A goal-directed spatial navigation model using forward trajectory planning based on grid cells: Forward linear look-ahead trajectory model. Eur. J. Neurosci. *35*, 916–931.
Etienne, A.S., and Jeffery, K.J. (2004). Path integration in mammals. Hippocampus *14*, 180–192.
Ezzyat, Y., and Davachi, L. (2014). Similarity Breeds Proximity: Pattern Similarity within and across Contexts Is Related to Later Mnemonic Judgments of Temporal Proximity. Neuron *81*, 1179–1189.

Felleman, D.J., Van Essen, D.C., and C, D. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. Cereb. Cortex *1*, 1–47.
Fernández, G., and Morris, R.G.M. (2018). Memory, Novelty and Prior Knowledge. Trends Neurosci. *41*, 654–659.
Fernández, G., Brewer, J.B., Zhao, Z., Glover, G.H., and Gabrieli, J.D.E. (1999). Level of sustained entorhinal activity at study correlates with subsequent cued-recall performance: A functional magnetic resonance imaging study with high acquisition rate. Hippocampus *9*, 35–44.
Fiete, I.R., Burak, Y., and Brookings, T. (2008). What Grid Cells Convey about Rat Location. J. Neurosci. *28*, 6858–6871.
Finkelstein, A., Derdikman, D., Rubin, A., Foerster, J.N., Las, L., and Ulanovsky, N. (2015). Three-dimensional head-direction coding in the bat brain. Nature *517*, 159–164.
Folkerts, S., Rutishauser, U., and Howard, M.W. (2018). Human episodic memory retrieval is accompanied by a neural contiguity effect. J. Neurosci. 2312–2317.
Foster, D.J. (2017). Replay Comes of Age. Annu. Rev. Neurosci. *40*, 581–602.
Foster, D., and Knierim, J. (2012). Sequence learning and the role of the hippocampus in rodent navigation. Curr. Opin. Neurobiol. *22*, 294–300.
Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440*, 680–683.
Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., and Frackowiak, R.S.J. (1994). Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapp. *2*, 189–210.
Fu, H., Rodriguez, G.A., Herman, M., Emrani, S., Nahmani, E., Barrett, G., Figueroa, H.Y., Goldberg, E., Hussaini, S.A., and Duff, K.E. (2017). Tau Pathology Induces Excitatory Neuron Loss, Grid Cell Dysfunction, and Spatial Memory Deficits Reminiscent of Early Alzheimer's Disease. Neuron *93*, 533-541.e5.
Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., and Moser, E.I. (2007). Hippocampal remapping and grid realignment in entorhinal cortex. Nature *446*, 190–194.
Fyhn, M., Hafting, T., Witter, M.P., Moser, E.I., and Moser, M.-B. (2008). Grid cells in mice. Hippocampus *18*, 1230–1238.
Gallistel, C.R. (1990). The organization of learning (Cambridge, Mass: MIT Press).
Gärdenfors, P. (2000). Conceptual spaces: the geometry of thought (Cambridge, Mass: MIT Press).
Gärdenfors, P., and Zenker, F. (2015). Editors' Introduction: Conceptual Spaces at Work. In Applications of Conceptual Spaces, (Springer, Cham), pp. 3–13.
Gardner, R.J., Lu, L., Wernle, T., Moser, M.-B., and Moser, E.I. (2017). Correlation structure of grid cells is preserved during sleep. BioRxiv 198499.
Garvert, M.M., Dolan, R.J., and Behrens, T.E. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. ELife *6*, e17086.
Gil, M., Ancau, M., Schlesiger, M.I., Neitz, A., Allen, K., De Marco, R.J., and Monyer, H. (2018). Impaired path integration in mice with disrupted grid cell firing. Nat. Neurosci. *21*, 81–91.
Gilboa, A., and Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. Trends Cogn. Sci. *21*, 618–631.
Gothard, K.M., Skaggs, W.E., Moore, K.M., and McNaughton, B.L. (1996a). Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task. J. Neurosci. *16*, 823–835.

Gothard, K.M., Skaggs, W.E., and McNaughton, B.L. (1996b). Dynamics of Mismatch Correction in the Hippocampal Ensemble Code for Space: Interaction between Path Integration and Environmental Cues. J. Neurosci. *16*, 8027–8040.

Grieves, R.M., and Jeffery, K.J. (2017). The representation of space in the brain. Behav. Processes *135*, 113–131.

Grosmark, A.D., and Buzsáki, G. (2016). Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. Science *351*, 1440–1443.

Gruber, M.J., Ritchey, M., Wang, S.-F., Doss, M.K., and Ranganath, C. (2016). Post-learning Hippocampal Dynamics Promote Preferential Retention of Rewarding Events. Neuron *0*.

Gupta, A.S., van der Meer, M. a a, Touretzky, D.S., and Redish, a D. (2010). Hippocampal replay is not a simple function of experience. Neuron *65*, 695–705.

Gustafson, N.J., and Daw, N.D. (2011). Grid Cells, Place Cells, and Geodesic Generalization for Spatial Reinforcement Learning. PLoS Comput. Biol. *7*.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E.I. (2005). Microstructure of a spatial map in the entorhinal cortex. Nature *436*, 801–806.

Hannula, D.E., and Duff, M.C. (2017). The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition (Cham: Springer International Publishing).

Hartley, T., Maguire, E.A., Spiers, H.J., and Burgess, N. (2003). The well-worn route and the path less traveled: distinct neural bases of route following and wayfinding in humans. Neuron *37*, 877–888.

Hartley, T., Trinkler, I., and Burgess, N. (2004). Geometric determinants of human spatial memory. Cognition *94*, 39–75.

Hartley, T., Bird, C.M., Chan, D., Cipolotti, L., Husain, M., Vargha-Khadem, F., and Burgess, N. (2007). The hippocampus is required for short-term topographical memory in humans. Hippocampus *17*, 34–48.

Hassabis, D., and Maguire, E.A. (2007). Deconstructing episodic memory with construction. Trends Cogn. Sci. *11*, 299–306.

Hassabis, D., Kumaran, D., Vann, S.D., and Maguire, E.A. (2007a). Patients with hippocampal amnesia cannot imagine new experiences. Proc. Natl. Acad. Sci. *104*, 1726–1731.

Hassabis, D., Kumaran, D., and Maguire, E.A. (2007b). Using Imagination to Understand the Neural Basis of Episodic Memory. J. Neurosci. *27*, 14365–14374.

Hassabis, D., Chu, C., Rees, G., Weiskopf, N., Molyneux, P.D., and Maguire, E. a (2009). Decoding neuronal ensembles in the human hippocampus. Curr. Biol. CB *19*, 546–554.

Hasselmo, M.E. (2009). A model of episodic memory: Mental time travel along encoded trajectories using grid cells. Neurobiol. Learn. Mem. *92*, 559–573.

Hasselmo, M.E. (2011). How We Remember: Brain Mechanisms of Episodic Memory (Cambridge, MA, US: The MIT Press).

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. Science *293*, 2425–2430.

Hayman, R., Verriotis, M.A., Jovalekic, A., Fenton, A.A., and Jeffery, K.J. (2011). Anisotropic encoding of three-dimensional space by place cells and grid cells. Nat. Neurosci. *14*, 1182.

Haynes, J.-D., and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. *8*, 686–691.

Heckers, S., Zalesak, M., Weiss, A.P., Ditman, T., and Titone, D. (2004). Hippocampal activation during transitive inference in humans. Hippocampus *14*, 153–162.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. Nat. Rev. Neurosci. *11*, 523–532.
Hermer, L., and Spelke, E.S. (1994). A geometric process for spatial reorientation in young children. Nature *370*, 57.
Herz, A.V., Mathis, A., and Stemmler, M. (2017). Periodic population codes: From a single circular variable to higher dimensions, multiple nested scales, and conceptual spaces. Curr. Opin. Neurobiol. *46*, 99–108.
Heys, J.G., and Dombeck, D.A. (2018). Evidence for a subcircuit in medial entorhinal cortex representing elapsed time during immobility. Nat. Neurosci. *21*, 1574.
Heys, J.G., Rangarajan, K.V., and Dombeck, D.A. (2014). The Functional Micro-organization of Grid Cells Revealed by Cellular-Resolution Imaging. Neuron *84*, 1079–1090.
Hok, V., Save, E., Lenck-Santini, P.P., and Poucet, B. (2005). Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. Proc. Natl. Acad. Sci. U. S. A. *102*, 4602–4607.
Hok, V., Lenck-Santini, P.-P., Roux, S., Save, E., Muller, R.U., and Poucet, B. (2007). Goal-Related Activity in Hippocampal Place Cells. J. Neurosci. *27*, 472–482.
Hollup, S.A., Molden, S., Donnett, J.G., Moser, M.-B., and Moser, E.I. (2001). Accumulation of Hippocampal Place Fields at the Goal Location in an Annular Watermaze Task. J. Neurosci. *21*, 1635–1644.
Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. *79*, 2554–2558.
Horner, A.J., Bisby, J.A., Zotow, E., Bush, D., and Burgess, N. (2016). Grid-like Processing of Imagined Navigation. Curr. Biol. *26*, 842–847.
Howard, M.W. (2018). Memory as Perception of the Past: Compressed Time in Mind and Brain. Trends Cogn. Sci. *22*, 124–136.
Howard, M.W., and Kahana, M.J. (2002). A Distributed Representation of Temporal Context. J. Math. Psychol. *46*, 269–299.
Howard, M.W., Fotedar, M.S., Datey, A.V., and Hasselmo, M.E. (2005). The Temporal Context Model in Spatial Navigation and Relational Learning: Toward a Common Explanation of Medial Temporal Lobe Function Across Domains. Psychol. Rev. *112*, 75–116.
Høydal, Ø.A., Skytøen, E.R., Moser, M.-B., and Moser, E.I. (2018). Object-vector coding in the medial entorhinal cortex. BioRxiv 286286.
Hsieh, L.-T., Gruber, M.J., Jenkins, L.J., and Ranganath, C. (2014). Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. Neuron *81*, 1165–1178.
Huettel, S.A., Song, A.W., and McCarthy, G. (2004). Functional magnetic resonance imaging (Sunderland, Mass: Sinauer Associates, Publishers).
Jacob, P.-Y., Casali, G., Spieser, L., Page, H., Overington, D., and Jeffery, K. (2017). An independent, landmark-dominated head-direction signal in dysgranular retrosplenial cortex. Nat. Neurosci. *20*, 173–175.
Jacobs, J., Kahana, M.J., Ekstrom, A.D., Mollison, M.V., and Fried, I. (2010). A sense of direction in human entorhinal cortex. Proc. Natl. Acad. Sci. U. S. A. *107*, 6487–6492.
Jacobs, J., Weidemann, C.T., Miller, J.F., Solway, A., Burke, J.F., Wei, X.-X., Suthana, N., Sperling, M.R., Sharan, A.D., Fried, I., et al. (2013). Direct recordings of grid-like neuronal activity in human spatial navigation. Nat. Neurosci. *16*, 1188–1190.
Jacobs, J., Miller, J., Lee, S.A., Coffey, T., Watrous, A.J., Sperling, M.R., Sharan, A., Worrell, G., Berry, B., Lega, B., et al. (2016). Direct Electrical Stimulation of the Human Entorhinal Region and Hippocampus Impairs Memory. Neuron *92*, 983–990.
Jadhav, S.P., Kemere, C., German, P.W., and Frank, L.M. (2012). Awake Hippocampal Sharp-Wave Ripples Support Spatial Memory. Science *336*, 1454–1458.

Jafarpour, A., and Spiers, H. (2017). Familiarity expands space and contracts time. Hippocampus *27*, 12–16.

Janzen, G., and van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. Nat. Neurosci. *7*, 673–677.

Javadi, A.-H., Emo, B., Howard, L.R., Zisch, F.E., Yu, Y., Knight, R., Silva, J.P., and Spiers, H.J. (2017). Hippocampal and prefrontal processing of network topology to simulate the future. Nat. Commun. *8*, 14652.

Jenkins, L.J., and Ranganath, C. (2010). Prefrontal and Medial Temporal Lobe Activity at Encoding Predicts Temporal Context Memory. J. Neurosci. *30*, 15558–15565.

Jenkins, L.J., and Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. Hippocampus *26*, 554–559.

Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. Med. Image Anal. *5*, 143–156.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. NeuroImage *17*, 825–841.

Johnson, A., and Redish, A.D. (2007). Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point. J. Neurosci. *27*, 12176–12189.

Julian, J.B., Fedorenko, E., Webster, J., and Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. NeuroImage *60*, 2357–2364.

Julian, J.B., Keinath, A.T., Frazzetta, G., and Epstein, R.A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. Nat. Neurosci. *21*, 191–194.

Jung, M.W., Wiener, S.I., and McNaughton, B.L. (1994). Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. J. Neurosci. *14*, 7347–7356.

Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. Nat. Neurosci. *8*, 679–685.

Kaplan, R., Bush, D., Bisby, J.A., Horner, A.J., Meyer, S.S., and Burgess, N. (2016). Medial Prefrontal–Medial Temporal Theta Phase Coupling in Dynamic Spatial Imagery. J. Cogn. Neurosci. *29*, 507–519.

Kaplan, R., Schuck, N.W., and Doeller, C.F. (2017a). The Role of Mental Maps in Decision-Making. Trends Neurosci. *40*, 256–259.

Kaplan, R., King, J., Koster, R., Penny, W.D., Burgess, N., and Friston, K.J. (2017b). The Neural Representation of Prospective Choice during Spatial Planning and Decisions. PLOS Biol. *15*, e1002588.

Karlsson, M.P., and Frank, L.M. (2009). Awake replay of remote experiences in the hippocampus. Nat. Neurosci. *12*, 913–918.

Kelly, J.W., McNamara, T.P., Bodenheimer, B., Carr, T.H., and Rieser, J.J. (2008). The shape of human navigation: How environmental geometry is used in maintenance of spatial orientation. Cognition *109*, 281–286.

van Kesteren, M.T.R., Fernández, G., Norris, D.G., and Hermans, E.J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. Proc. Natl. Acad. Sci. *107*, 7550–7555.

van Kesteren, M.T.R., Ruiter, D.J., Fernández, G., and Henson, R.N. (2012). How schema and novelty augment memory formation. Trends Neurosci. *35*, 211–219.

van Kesteren, M.T.R., Rijpkema, M., Ruiter, D.J., Morris, R.G.M., and Fernández, G. (2014). Building on Prior Knowledge: Schema-dependent Encoding Processes Relate to Academic Performance. J. Cogn. Neurosci. *26*, 2250–2261.

Killian, N.J., Jutras, M.J., and Buffalo, E.A. (2012). A map of visual space in the primate entorhinal cortex. Nature *491*, 761–764.
Killian, N.J., Potter, S.M., and Buffalo, E.A. (2015). Saccade direction encoding in the primate entorhinal cortex during visual exploration. Proc. Natl. Acad. Sci. 201417059.
Kim, M., and Maguire, E.A. (2018). 3D grid cells in human entorhinal cortex: Theoretical and methodological considerations and fMRI findings. BioRxiv 282327.
Kim, M., Jeffery, K.J., and Maguire, E.A. (2017). Multivoxel Pattern Analysis Reveals 3D Place Information in the Human Hippocampus. J. Neurosci. *37*, 4270–4279.
Kjelstrup, K.B., Solstad, T., Brun, V.H., Hafting, T., Leutgeb, S., Witter, M.P., Moser, E.I., and Moser, M.-B. (2008). Finite Scale of Spatial Representation in the Hippocampus. Science *321*, 140–143.
Knierim, J.J., Neunuebel, J.P., and Deshmukh, S.S. (2014). Functional correlates of the lateral and medial entorhinal cortex: objects, path integration and local–global reference frames. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130369.
Komorowski, R.W., Garcia, C.G., Wilson, A., Hattori, S., Howard, M.W., and Eichenbaum, H. (2013). Ventral Hippocampal Neurons Are Shaped by Experience to Represent Behaviorally Relevant Contexts. J. Neurosci. *33*, 8079–8087.
Konkel, A., and Cohen, N.J. (2009). Relational memory and the hippocampus: Representations and methods. Front. Neurosci. *3*.
Kosslyn, S.M., Ball, T.M., and Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. J. Exp. Psychol. Hum. Percept. Perform. *4*, 47–60.
Koster, R., Chadwick, M.J., Chen, Y., Berron, D., Banino, A., Düzel, E., Hassabis, D., and Kumaran, D. (2018). Big-Loop Recurrence within the Hippocampal System Supports Integration of Information across Episodes. Neuron *99*, 1342-1354.e6.
Kraus, B.J., Robinson II, R.J., White, J.A., Eichenbaum, H., and Hasselmo, M.E. (2013). Hippocampal "Time Cells": Time versus Path Integration. Neuron *78*, 1090–1101.
Kraus, B.J., Brandon, M.P., Robinson, R.J., Connerney, M.A., Hasselmo, M.E., and Eichenbaum, H. (2015). During Running in Place, Grid Cells Integrate Elapsed Time and Distance Run. Neuron *88*, 578–589.
Kravitz, D.J., Peng, C.S., and Baker, C.I. (2011). Real-World Scene Representations in High-Level Visual Cortex: It's the Spaces More Than the Places. J. Neurosci. *31*, 7322–7333.
Kreiman, G., Koch, C., and Fried, I. (2000). Imagery neurons in the human brain. Nature *408*, 357–361.
Kriegeskorte, N., and Kievit, R.A. (2013). Representational geometry: integrating cognition, computation, and the brain. Trends Cogn. Sci. *17*, 401–412.
Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. Proc. Natl. Acad. Sci. *103*, 3863–3868.
Kriegeskorte, N., Mur, M., and Bandettini, P. (2008a). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4–4.
Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P.A. (2008b). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. Neuron *60*, 1126–1141.
Kropff, E., Carmichael, J.E., Moser, M.-B., and Moser, E.I. (2015). Speed cells in the medial entorhinal cortex. Nature *523*, 419–424.
Krupic, J., Bauza, M., Burton, S., Barry, C., and O'Keefe, J. (2015). Grid cell symmetry is shaped by environmental geometry. Nature *518*, 232–235.
Krupic, J., Bauza, M., Burton, S., and O'Keefe, J. (2018). Local transformations of the hippocampal cognitive map. Science *359*, 1143–1146.

Kubie, J.L., and Fenton, A.A. (2012). Linear Look-Ahead in Conjunctive Cells: An Entorhinal Mechanism for Vector-Based Navigation. Front. Neural Circuits *6*.

Kudrimoti, H.S., Barnes, C.A., and McNaughton, B.L. (1999). Reactivation of Hippocampal Cell Assemblies: Effects of Behavioral State, Experience, and EEG Dynamics. J. Neurosci. *19*, 4090–4101.

Kuhl, B.A., Shah, A.T., DuBrow, S., and Wagner, A.D. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. Nat. Neurosci. *13*, 501–506.

Kumaran, D., and Maguire, E.A. (2006). An Unexpected Sequence of Events: Mismatch Detection in the Human Hippocampus. PLOS Biol. *4*, e424.

Kumaran, D., and McClelland, J.L. (2012). Generalization Through the Recurrent Interaction of Episodic Memories. Psychol. Rev. *119*, 573–616.

Kumaran, D., Melo, H.L., and Duzel, E. (2012). The Emergence and Representation of Knowledge about Social and Nonsocial Hierarchies. Neuron *76*, 653–666.

Kumaran, D., Banino, A., Blundell, C., Hassabis, D., and Dayan, P. (2016a). Computations Underlying Social Hierarchy Learning: Distinct Neural Mechanisms for Updating and Representing Self-Relevant Information. Neuron *92*, 1135–1147.

Kumaran, D., Hassabis, D., and McClelland, J.L. (2016b). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. Trends Cogn. Sci. *20*, 512–534.

Kunz, L., Navarro Schröder, T., Lee, H., Montag, C., Lachmann, B., Sariyska, R., Reuter, M., Stirnberg, R., Stöcker, T., Messing-Floeter, P.C., et al. (2015). Reduced grid-cell–like representations in adults at genetic risk for Alzheimer's disease. Science *350*, 430–433.

Kurth-Nelson, Z., Economides, M., Dolan, R.J., and Dayan, P. (2016). Fast Sequences of Non-spatial State Representations in Humans. Neuron *91*, 194–204.

Kyle, C.T., Smuda, D.N., Hassan, A.S., and Ekstrom, A.D. (2015). Roles of human hippocampal subfields in retrieval of spatial and temporal context. Behav. Brain Res. *278*, 549–558.

Lee, A.K., and Wilson, M.A. (2002). Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep. Neuron *36*, 1183–1194.

Leutgeb, J.K., Leutgeb, S., Treves, A., Meyer, R., Barnes, C.A., McNaughton, B.L., Moser, M.-B., and Moser, E.I. (2005). Progressive Transformation of Hippocampal Neuronal Representations in "Morphed" Environments. Neuron *48*, 345–358.

Leutgeb, S., Leutgeb, J.K., Treves, A., Moser, M.-B., and Moser, E.I. (2004). Distinct Ensemble Codes in Hippocampal Areas CA3 and CA1. Science *305*, 1295–1298.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., and Burgess, N. (2009). Boundary Vector Cells in the Subiculum of the Hippocampal Formation. J. Neurosci. *29*, 9771–9777.

Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. Nature *453*, 869–878.

Lositsky, O., Chen, J., Toker, D., Honey, C.J., Shvartsman, M., Poppenk, J.L., Hasson, U., and Norman, K.A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. ELife *5*, e16070.

Maass, A., Berron, D., Libby, L., Ranganath, C., and Düzel, E. (2015). Functional subregions of the human entorhinal cortex. ELife *4*, e06426.

MacDonald, C.J., Lepage, K.Q., Eden, U.T., and Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. Neuron *71*, 737–749.

Mack, M.L., Love, B.C., and Preston, A.R. (2017). Building concepts one episode at a time: The hippocampus and concept formation. Neurosci. Lett.

Maguire, E.A., Burgess, N., Donnett, James G., Frackowiak, R.S., Frith, C.D., and O'Keefe, J. (1998). Knowing Where and Getting There: A Human Navigation Network. Science *280*, 921–924.
Maguire, E.A., Vargha-Khadem, F., and Mishkin, M. (2001). The effects of bilateral hippocampal damage on fMRI regional activations and interactions during memory retrieval. Brain *124*, 1156–1170.
Maidenbaum, S., Miller, J., Stein, J.M., and Jacobs, J. (2018). Grid-like hexadirectional modulation of human entorhinal theta oscillations. Proc. Natl. Acad. Sci. *115*, 10798–10803.
Marchette, S.A., Vass, L.K., Ryan, J., and Epstein, R.A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. Nat. Neurosci. *17*, 1598–1606.
Marchette, S.A., Vass, L.K., Ryan, J., and Epstein, R.A. (2015). Outside Looking In: Landmark Generalization in the Human Navigational System. J. Neurosci. *35*, 14896–14908.
Marr, D. (1971). Simple memory: a theory for archicortex. Phil Trans R Soc Lond B *262*, 23–81.
Mathis, A., Herz, A.V.M., and Stemmler, M. (2012). Optimal Population Codes for Space: Grid Cells Outperform Place Cells. Neural Comput. *24*, 2280–2317.
Mau, W., Sullivan, D.W., Kinsky, N.R., Hasselmo, M.E., Howard, M.W., and Eichenbaum, H. (2018). The Same Hippocampal CA1 Population Simultaneously Codes Temporal Information over Multiple Timescales. Curr. Biol. *28*, 1499-1508.e4.
McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights From the Successes and Failures of Connectionist Models of Learning and Memory. Psychol. Rev. *102*, 419–457.
McKenzie, S., Frank, A.J., Kinsky, N.R., Porter, B., Rivière, P.D., and Eichenbaum, H. (2014). Hippocampal Representation of Related and Opposing Memories Develop within Distinct, Hierarchically Organized Neural Schemas. Neuron *83*, 202–215.
McNaughton, B.L., and Morris, R.G.M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends Neurosci. *10*, 408–415.
McNaughton, B.L., Battaglia, F.P., Jensen, O., Moser, E.I., and Moser, M.-B. (2006). Path integration and the neural basis of the "cognitive map". Nat. Rev. Neurosci. *7*, 663–678.
van der Meer, M.A., and Redish, A.D. (2009). Covert expectation-of-reward in rat ventral striatum at decision points. Front. Integr. Neurosci. *3*.
van der Meer, M.A.A., Johnson, A., Schmitzer-Torbert, N.C., and Redish, A.D. (2010). Triple Dissociation of Information Processing in Dorsal Striatum, Ventral Striatum, and Hippocampus on a Learned Spatial Decision Task. Neuron *67*, 25–32.
Milivojevic, B., and Doeller, C.F. (2013). Mnemonic Networks in the Hippocampal Formation: From Spatial Maps to Temporal and Conceptual Codes. J. Exp. Psychol. Gen.
Milivojevic, B., Vicente-Grabovetsky, A., and Doeller, C.F. (2015). Insight Reconfigures Hippocampal-Prefrontal Memories. Curr. Biol. *25*, 821–830.
Miller, J.F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow, M., Polyn, S.M., Jacobs, J., et al. (2013). Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. Science *342*, 1111–1114.
Milner, B. (1962). Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. In Physiologie de l'Hippocampe, P. Passouant, ed. (Paris: Éditions Recherche Scientifique), pp. 257–272.
Mittelstaedt, M.-L., and Mittelstaedt, H. (1980). Homing by path integration in a mammal. Naturwissenschaften *67*, 566–567.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. Nature *518*, 529–533.

Montchal, M.E., Reagh, Z.M., and Yassa, M.A. (2019). Precise temporal memories are supported by the lateral entorhinal cortex in humans. Nat. Neurosci. 1.

Moscovitch, M., Cabeza, R., Winocur, G., and Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. Annu. Rev. Psychol. *67*, 105–134.

Moser, E.I., Kropff, E., and Moser, M.-B. (2008). Place Cells, Grid Cells, and the Brain's Spatial Representation System. Annu. Rev. Neurosci. *31*, 69–89.

Moser, E.I., Roudi, Y., Witter, M.P., Kentros, C., Bonhoeffer, T., and Moser, M.-B. (2014). Grid cells and cortical representation. Nat. Rev. Neurosci. *15*, 466–481.

Moser, E.I., Moser, M.-B., and McNaughton, B.L. (2017). Spatial representation in the hippocampal formation: a history. Nat. Neurosci. *20*, 1448–1464.

Muller, R.U., and Kubie, J.L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J. Neurosci. *7*, 1951–1968.

Nadasdy, Z., Nguyen, T.P., Török, Á., Shen, J.Y., Briggs, D.E., Modur, P.N., and Buchanan, R.J. (2017). Context-dependent spatially periodic activity in the human entorhinal cortex. Proc. Natl. Acad. Sci. *114*, E3516–E3525.

Nau, M., Navarro Schröder, T., Bellmund, J.L.S., and Doeller, C.F. (2018a). Hexadirectional coding of visual space in human entorhinal cortex. Nat. Neurosci. *21*, 188–190.

Nau, M., Julian, J.B., and Doeller, C.F. (2018b). How the Brain's Navigation System Shapes Our Visual Experience. Trends Cogn. Sci. *22*, 810–825.

Navarro Schröder, T., Haak, K.V., Zaragoza Jimenez, N.I., Beckmann, C.F., and Doeller, C.F. (2015). Functional topography of the human entorhinal cortex. ELife *4*, e06738.

Navarro Schröder, T., Towse, B.W., Burgess, N., Barry, C., and Doeller, C.F. (2017). Optimal decision making using grid cells under spatial uncertainty. BioRxiv 166306.

Nielson, D.M., Smith, T.A., Sreekumar, V., Dennis, S., and Sederberg, P.B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. Proc. Natl. Acad. Sci. *112*, 11078–11083.

Nolan, C.R., Vromen, J.M.G., Cheung, A., and Baumann, O. (2018). Evidence against the Detectability of a Hippocampal Place Code Using Functional Magnetic Resonance Imaging. ENeuro *5*, ENEURO.0177-18.2018.

Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. *10*, 424–430.

O'Keefe, J., and Conway, D.H. (1978). Hippocampal place units in the freely moving rat: Why they fire where they fire. Exp. Brain Res. *31*, 573–590.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. *34*, 171–175.

O'Keefe, J., and Nadel, L. (1978). The Hippocampus as a Cognitive Map (Oxford: Clarendon Press).

O'Keefe, J., and Nadel, L. (1979). Précis of O'Keefe & Nadel's The hippocampus as a cognitive map. Behav. Brain Sci. *2*, 487–494.

O'Keefe, J., Burgess, N., Donnett, J.G., Jeffery, K.J., and Maguire, E.A. (1998). Place cells, navigational accuracy, and the human hippocampus. Philos. Trans. R. Soc. B Biol. Sci. *353*, 1333–1340.

Ólafsdóttir, H.F., Barry, C., Saleem, A.B., Hassabis, D., and Spiers, H.J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. ELife *4*, e06063.

Ólafsdóttir, H.F., Carpenter, F., and Barry, C. (2016). Coordinated grid and place cell replay during rest. Nat. Neurosci. *19*, 792–794.
Ólafsdóttir, H.F., Carpenter, F., and Barry, C. (2017). Task Demands Predict a Dynamic Switch in the Content of Awake Hippocampal Replay. Neuron *96*, 925-935.e6.
Ólafsdóttir, H.F., Bush, D., and Barry, C. (2018). The Role of Hippocampal Replay in Memory and Planning. Curr. Biol. *28*, R37–R50.
Omer, D.B., Maimon, S.R., Las, L., and Ulanovsky, N. (2018). Social place-cells in the bat hippocampus. Science *359*, 218–224.
O'Neill, J., Senior, T., and Csicsvari, J. (2006). Place-Selective Firing of CA1 Pyramidal Cells during Sharp Wave/Ripple Network Patterns in Exploratory Behavior. Neuron *49*, 143–155.
O'Neill, J., Boccara, C.N., Stella, F., Schoenenberger, P., and Csicsvari, J. (2017). Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. Science *355*, 184–188.
Paller, K.A., and Wagner, A.D. (2002). Observing the transformation of experience into memory. Trends Cogn. Sci. *6*, 93–102.
Park, S., Brady, T.F., Greene, M.R., and Oliva, A. (2011). Disentangling Scene Content from Spatial Boundary: Complementary Roles for the Parahippocampal Place Area and Lateral Occipital Complex in Representing Real-World Scenes. J. Neurosci. *31*, 1333–1340.
Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally Generated Cell Assembly Sequences in the Rat Hippocampus. Science *321*, 1322–1327.
Pernet, C.R., Wilcox, R.R., and Rousselet, G.A. (2013). Robust Correlation Analyses: False Positive and Power Validation Using a New Open Source Matlab Toolbox. Front. Psychol. *3*.
Peters, J., and Büchel, C. (2010). Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions. Neuron *66*, 138–148.
Peyrache, A., Lacroix, M.M., Petersen, P.C., and Buzsáki, G. (2015). Internally organized mechanisms of the head direction sense. Nat. Neurosci. *18*, 569–575.
Pfeiffer, B.E., and Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. Nature *497*, 74–79.
Pfeiffer, B.E., and Foster, D.J. (2015). Autoassociative dynamics in the generation of sequences of hippocampal place cells. Science *349*, 180–183.
Polyn, S.M., Natu, V.S., Cohen, J.D., and Norman, K.A. (2005). Category-Specific Cortical Activity Precedes Retrieval During Memory Search. Science *310*, 1963–1966.
Poulter, S., Hartley, T., and Lever, C. (2018). The Neurobiology of Mammalian Navigation. Curr. Biol. *28*, R1023–R1042.
Preston, A.R., and Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. Curr. Biol. *23*, R764-73.
Preston, A.R., Shrager, Y., Dudukovic, N.M., and Gabrieli, J.D.E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus *14*, 148–152.
Pylyshyn, Z.W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. Psychol. Bull. *80*, 1–24.
Qasim, S.E., Miller, J., Inman, C.S., Gross, R., Willie, J.T., Lega, B., Lin, J.-J., Sharan, A., Wu, C., Sperling, M.R., et al. (2018). Single neurons in the human entorhinal cortex remap to distinguish individual spatial memories. BioRxiv 433862.
Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. Nature *435*, 1102–1107.

Radvansky, B.A., and Dombeck, D.A. (2018). An olfactory virtual reality system for mice. Nat. Commun. *9*, 839.

Ranganath, C. (2018). Time, memory, and the legacy of Howard Eichenbaum. Hippocampus 1–16.

Ranganath, C., and Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. Nat. Rev. Neurosci. *13*, 713–726.

Redish, A.D. (2016). Vicarious trial and error. Nat. Rev. Neurosci. *17*, 147–159.

Redish, A.D., Battaglia, F.P., Chawla, M.K., Ekstrom, A.D., Gerrard, J.L., Lipa, P., Rosenzweig, E.S., Worley, P.F., Guzowski, J.F., McNaughton, B.L., et al. (2001). Independence of Firing Correlates of Anatomically Proximate Hippocampal Pyramidal Cells. J. Neurosci. *21*, RC134–RC134.

Ritchey, M., Libby, L.A., and Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition: the PMAT framework. In Progress in Brain Research, (Elsevier), pp. 45–64.

Rodriguez, P.F. (2010). Neural decoding of goal locations in spatial navigation in humans with fMRI. Hum. Brain Mapp. *31*, 391–397.

Rosch, E. (1975). Cognitive reference points. Cognit. Psychol. *7*, 532–547.

Rotenberg, A., Mayford, M., Hawkins, R.D., Kandel, E.R., and Muller, R.U. (1996). Mice Expressing Activated CaMKII Lack Low Frequency LTP and Do Not Form Stable Place Cells in the CA1 Region of the Hippocampus. Cell *87*, 1351–1361.

Rowland, D.C., Roudi, Y., Moser, M.-B., and Moser, E.I. (2016). Ten Years of Grid Cells. Annu. Rev. Neurosci. *39*, null.

Sarel, A., Finkelstein, A., Las, L., and Ulanovsky, N. (2017). Vectorial representation of spatial goals in the hippocampus of bats. Science *355*, 176–180.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.-B., and Moser, E.I. (2006). Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex. Science *312*, 758–762.

Savelli, F., Yoganarasimha, D., and Knierim, J.J. (2008). Influence of boundary removal on the spatial representations of the medial entorhinal cortex. Hippocampus *18*, 1270–1282.

Schacter, D.L., and Wagner, A.D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. Hippocampus *9*, 7–24.

Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. *8*, 657–661.

Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., and Szpunar, K.K. (2012). The Future of Memory: Remembering, Imagining, and the Brain. Neuron *76*, 677–694.

Schapiro, A.C., Kustner, L.V., and Turk-Browne, N.B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. Curr. Biol. CB *22*, 1622–1627.

Schapiro, A.C., Turk-Browne, N.B., Norman, K.A., and Botvinick, M.M. (2016). Statistical learning of temporal community structure in the hippocampus. Hippocampus *26*, 3–8.

Schapiro, A.C., McDevitt, E.A., Rogers, T.T., Mednick, S.C., and Norman, K.A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. Nat. Commun. *9*, 3920.

Schiller, D., Eichenbaum, H., Buffalo, E.A., Davachi, L., Foster, D.J., Leutgeb, S., and Ranganath, C. (2015). Memory and Space: Towards an Understanding of the Cognitive Map. J. Neurosci. *35*, 13904–13911.

Schinazi, V.R., Nardi, D., Newcombe, N.S., Shipley, T.F., and Epstein, R.A. (2013). Hippocampal size predicts rapid learning of a cognitive map in humans. Hippocampus *23*, 515–528.
Schlichting, M.L., and Preston, A.R. (2014). Memory reactivation during rest supports upcoming learning of related content. Proc. Natl. Acad. Sci. 201404396.
Schlichting, M.L., and Preston, A.R. (2015). Memory integration: neural mechanisms and implications for behavior. Curr. Opin. Behav. Sci. *1*, 1–8.
Schlichting, M.L., Mumford, J.A., and Preston, A.R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat. Commun. *6*.
Schuck, N.W., and Niv, Y. (2018). Sequential replay of non-spatial task states in the human hippocampus. BioRxiv 315978.
Schuck, N.W., Doeller, C.F., Polk, T.A., Lindenberger, U., and Li, S.-C. (2015a). Human aging alters the neural computation and representation of space. NeuroImage *117*, 141–150.
Schuck, N.W., Gaschler, R., Wenke, D., Heinzle, J., Frensch, P.A., Haynes, J.-D., and Reverberi, C. (2015b). Medial Prefrontal Cortex Predicts Internally Driven Strategy Shifts. Neuron *86*, 331–340.
Schuck, N.W., Cai, M.B., Wilson, R.C., and Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. Neuron *91*, 1402–1412.
Scoville, W.B., and Milner, B. (1957). Loss of Recent Memory After Bilateral Hippocampal Lesions. J. Neurol. Neurosurg. Psychiatry *20*, 11–21.
Sharpe, M.J., Stalnaker, T., Schuck, N.W., Killcross, S., Schoenbaum, G., and Niv, Y. (2019). An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. Annu. Rev. Psychol. *70*, null.
Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. Science *237*, 1317–1323.
Shine, J.P., Valdés-Herrera, J.P., Hegarty, M., and Wolbers, T. (2016). The Human Retrosplenial Cortex and Thalamus Code Head Direction in a Global Reference Frame. J. Neurosci. *36*, 6371–6381.
Shohamy, D., and Wagner, A.D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. Neuron *60*, 378–389.
Singer, A.C., Carr, M.F., Karlsson, M.P., and Frank, L.M. (2013). Hippocampal SWR Activity Predicts Correct Decisions during the Initial Learning of an Alternation Task. Neuron *77*, 1163–1173.
Skaggs, W.E., and McNaughton, B.L. (1996). Replay of Neuronal Firing Sequences in Rat Hippocampus During Sleep Following Spatial Experience. Science *271*, 1870–1873.
Solstad, T., Boccara, C.N., Kropff, E., Moser, M.-B., and Moser, E.I. (2008). Representation of Geometric Borders in the Entorhinal Cortex. Science *322*, 1865–1868.
Spelke, E.S., and Lee, S.A. (2012). Core systems of geometry in animal minds. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 2784–2793.
Spiers, H.J., and Maguire, E. a (2006). Thoughts, behaviour, and brain dynamics during navigation in the real world. NeuroImage *31*, 1826–1840.
Spiers, H.J., and Maguire, E.A. (2007). A Navigational Guidance System in the Human Brain. *626*, 618–626.
Spreng, R.N., Mar, R.A., and Kim, A.S.N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J. Cogn. Neurosci. *21*, 489–510.
Squire, L.R. (1982). The Neuropsychology of Human Memory. Annu. Rev. Neurosci. *5*, 241–273.

Squire, L.R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. Psychol. Rev. *99*, 195–231.
Squire, L.R. (2009). The Legacy of Patient H.M. for Neuroscience. Neuron *61*, 6–9.
Squire, L.R., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. Science *253*, 1380–1386.
Stachenfeld, K.L., Botvinick, M.M., and Gershman, S.J. (2017). The hippocampus as a predictive map. Nat. Neurosci. *20*, 1643–1653.
Stangl, M., Shine, J., and Wolbers, T. (2017). The GridCAT: A Toolbox for Automated Analysis of Human Grid Cell Codes in fMRI. Front. Neuroinformatics *11*.
Stangl, M., Achtzehn, J., Huber, K., Dietrich, C., Tempelmann, C., and Wolbers, T. (2018). Compromised Grid-Cell-like Representations in Old Age as a Key Mechanism to Explain Age-Related Navigational Deficits. Curr. Biol. *28*, 1108-1115.e6.
Staresina, B.P., Alink, A., Kriegeskorte, N., and Henson, R.N. (2013). Awake reactivation predicts memory in humans. Proc. Natl. Acad. Sci. *110*, 21159–21164.
Staudigl, T., Leszczynski, M., Jacobs, J., Sheth, S.A., Schroeder, C.E., Jensen, O., and Doeller, C.F. (2018). Hexadirectional Modulation of High-Frequency Electrophysiological Activity in the Human Anterior Medial Temporal Lobe Maps Visual Space. Curr. Biol. *28*, 3325-3329.e4.
Steemers, B., Vicente-Grabovetsky, A., Barry, C., Smulders, P., Navarro Schröder, T., Burgess, N., and Doeller, C.F. (2016). Hippocampal Attractor Dynamics Predict Memory-Based Decision Making. Curr. Biol. *26*, 1750–1757.
Stelzer, J., Chen, Y., and Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. NeuroImage *65*, 69–82.
Stemmler, M., Mathis, A., and Herz, A.V.M. (2015). Connecting multiple spatial scales to decode the population activity of grid cells. Sci. Adv. *1*, e1500816.
Stensola, H., Stensola, T., Solstad, T., Frøland, K., Moser, M.-B., and Moser, E.I. (2012). The entorhinal grid map is discretized. Nature *492*, 72–78.
Stensola, T., Stensola, H., Moser, M.-B., and Moser, E.I. (2015). Shearing-induced asymmetry in entorhinal grid cells. Nature *518*, 207–212.
Stokes, J., Kyle, C., and Ekstrom, A.D. (2015). Complementary Roles of Human Hippocampal Subfields in Differentiation and Integration of Spatial Context. J. Cogn. Neurosci. *27*, 546–559.
Stolk, A., Verhagen, L., and Toni, I. (2016). Conceptual Alignment: How Brains Achieve Mutual Understanding. Trends Cogn. Sci. *20*, 180–191.
Strange, B.A., Witter, M.P., Lein, E.S., and Moser, E.I. (2014). Functional organization of the hippocampal longitudinal axis. Nat. Rev. Neurosci. *15*, 655–669.
van Strien, N.M., Cappaert, N.L.M., and Witter, M.P. (2009). The anatomy of memory: an interactive overview of the parahippocampal–hippocampal network. Nat. Rev. Neurosci. *10*, 272–282.
Sulpizio, V., Committeri, G., and Galati, G. (2014). Distributed cognitive maps reflecting real distances between places and views in the human brain. Front. Hum. Neurosci. *8*, 716.
Sun, C., Kitamura, T., Yamamoto, J., Martin, J., Pignatelli, M., Kitch, L.J., Schnitzer, M.J., and Tonegawa, S. (2015). Distinct speed dependence of entorhinal island and ocean cells, including respective grid cells. Proc. Natl. Acad. Sci. *112*, 9466–9471.
Tambini, A., and Davachi, L. (2013). Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. Proc. Natl. Acad. Sci. *110*, 19591–19596.
Taube, J.S. (2007). The head direction signal: origins and sensory-motor integration. Annu. Rev. Neurosci. *30*, 181–207.

Taube, J.S., Muller, R.U., and Ranck, J.B. (1990a). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. J. Neurosci. *10*, 420–435.
Taube, J.S., Muller, R.U., and Ranck, J.B. (1990b). Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. J. Neurosci. *10*, 436–447.
Taube, J.S., Valerio, S., and Yoder, R.M. (2013). Is Navigation in Virtual Reality with fMRI Really Navigation? J. Cogn. Neurosci. *25*, 1008–1019.
Tavares, R.M., Mendelsohn, A., Grossman, Y., Williams, C.H., Shapiro, M., Trope, Y., and Schiller, D. (2015). A Map for Social Navigation in the Human Brain. Neuron *87*, 231–243.
Tcheang, L., Bülthoff, H.H., and Burgess, N. (2011). Visual influence on path integration in darkness indicates a multimodal representation of large-scale space. Proc. Natl. Acad. Sci. *108*, 1152–1157.
Tenenbaum, J.B., Kemp, C., Griffiths, T.L., and Goodman, N.D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. Science *331*, 1279–1285.
Thavabalasingam, S., O'Neil, E.B., and Lee, A.C.H. (2018). Multivoxel pattern similarity suggests the integration of temporal duration in hippocampal event sequence representations. NeuroImage *178*, 136–146.
Tolman, E.C. (1948). Cognitive maps in rats and men. Psychol. Rev. *55*, 189–208.
Trettel, S.G., Trimper, J.B., Hwaun, E., Fiete, I.R., and Colgin, L.L. (2017). Grid cell co-activity patterns during sleep reflect spatial overlap of grid fields during active behaviors. BioRxiv 198671.
Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J.J., Moser, M.-B., and Moser, E.I. (in press). Integrating time from experience in lateral entorhinal cortex. Nature.
Tsao, A., Moser, M.-B., and Moser, E.I. (2013). Traces of Experience in the Lateral Entorhinal Cortex. Curr. Biol. *23*, 399–405.
Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J.J., Moser, M.-B., and Moser, E.I. (2018). Integrating time from experience in the lateral entorhinal cortex. Nature *561*, 57–62.
Tubridy, S., and Davachi, L. (2011). Medial Temporal Lobe Contributions to Episodic Sequence Encoding. Cereb. Cortex *21*, 272–280.
Tulving, E. (1972). Episodic and semantic memory. In Organization of Memory, E. Tulving, and W. Donaldson, eds. (Oxford, England: Academic Press), pp. 381–403.
Tulving, E. (1983). Elements of Episodic Memory (Clarendon Press).
Tulving, E. (1985). Elements of Episodic Memory (Oxford, New York: Oxford University Press).
Tulving, E. (2002). Episodic Memory: From Mind to Brain. Annu. Rev. Psychol. *53*, 1–25.
Ulanovsky, N., and Moss, C.F. (2007). Hippocampal cellular and network activity in freely moving echolocating bats. Nat. Neurosci. *10*, 224–233.
Vann, S.D., Aggleton, J.P., and Maguire, E. a (2009). What does the retrosplenial cortex do? Nat. Rev. Neurosci. *10*, 792–802.
Vass, L.K., and Epstein, R.A. (2013). Abstract representations of location and facing direction in the human brain. J. Neurosci. *33*, 6133–6142.
Vass, L.K., and Epstein, R.A. (2016). Common Neural Representations for Visually Guided Reorientation and Spatial Imagery. Cereb. Cortex bhv343.
Vickerstaff, R.J., and Cheung, A. (2010). Which coordinate system for modelling path integration? J. Theor. Biol. *263*, 242–261.

Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R., and Buckner, R.L. (1998). Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity. Science *281*, 1188–1191.

Wagner, I.C., van Buuren, M., Kroes, M.C., Gutteling, T.P., van der Linden, M., Morris, R.G., and Fernández, G. (2015). Schematic memory components converge within angular gyrus during retrieval. ELife *4*, e09668.

Wang, F., and Diana, R.A. (2017). Temporal context in human fMRI. Curr. Opin. Behav. Sci. *17*, 57–64.

Warren, W.H., Rothman, D.B., Schnapp, B.H., and Ericson, J.D. (2017). Wormholes in virtual space: From cognitive maps to cognitive graphs. Cognition *166*, 152–163.

Whittington, J.C.R., Muller, T.H., Barry, C., and Behrens, T.E.J. (2018). Generalisation of structural knowledge in the Hippocampal-Entorhinal system. ArXiv180509042 Cs Q-Bio Stat.

Wikenheiser, A.M., and Schoenbaum, G. (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. Nat. Rev. Neurosci. *17*, 513–523.

Wills, T.J., Lever, C., Cacucci, F., Burgess, N., and O'Keefe, J. (2005). Attractor Dynamics in the Hippocampal Representation of the Local Environment. Science *308*, 873–876.

Wilming, N., König, P., König, S., and Buffalo, E.A. (2018). Entorhinal cortex receptive fields are modulated by spatial attention, even without movement. ELife *7*, e31745.

Wilson, M.A., and McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. Science *265*, 676–679.

Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., and Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. Neuron *81*, 267–279.

Witter, M.P., Kleven, H., and Flatmoen, A.K. (2017a). Comparative Contemplations on the Hippocampus. Brain. Behav. Evol. *90*, 15–24.

Witter, M.P., Doan, T.P., Jacobsen, B., Nilssen, E.S., and Ohara, S. (2017b). Architecture of the Entorhinal Cortex A Review of Entorhinal Anatomy in Rodents with Some Comparative Notes. Front. Syst. Neurosci. *11*.

Wolbers, T., and Büchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. J. Neurosci. Off. J. Soc. Neurosci. *25*, 3333–3340.

Wolbers, T., Weiller, C., and Büchel, C. (2004). Neural foundations of emerging route knowledge in complex spatial environments. Brain Res. Cogn. Brain Res. *21*, 401–411.

Wolbers, T., Wiener, J.M., Mallot, H.A., and Büchel, C. (2007). Differential recruitment of the hippocampus, medial prefrontal cortex, and the human motion complex during path integration in humans. J. Neurosci. *27*, 9408–9416.

Wolbers, T., Klatzky, R.L., Loomis, J.M., Wutte, M.G., and Giudice, N.A. (2011). Modality-Independent Coding of Spatial Layout in the Human Brain. Curr. Biol. *21*, 984–989.

Wood, E.R., Dudchenko, P.A., and Eichenbaum, H. (1999). The global record of memory in hippocampal neuronal activity. Nature *397*, 613–616.

Wood, E.R., Dudchenko, P.A., Robitsek, R.J., and Eichenbaum, H. (2000). Hippocampal Neurons Encode Information about Different Types of Memory Episodes Occurring in the Same Location. Neuron *27*, 623–633.

Wu, C.-T., Haggerty, D., Kemere, C., and Ji, D. (2017). Hippocampal awake replay in fear memory retrieval. Nat. Neurosci. *20*, 571–580.

Yartsev, M.M., and Ulanovsky, N. (2013). Representation of Three-Dimensional Space in the Hippocampus of Flying Bats. Science *340*, 367–372.

Yartsev, M.M., Witter, M.P., and Ulanovsky, N. (2011). Grid cells without theta oscillations in the entorhinal cortex of bats. Nature *479*, 103–107.

Yoon, K., Buice, M.A., Barry, C., Hayman, R., Burgess, N., and Fiete, I.R. (2013). Specific evidence of low-dimensional continuous attractor dynamics in grid cells. Nat. Neurosci. *16*, 1077–1084.

Young, B.J., Fox, G.D., and Eichenbaum, H. (1994). Correlates of hippocampal complex-spike cell activity in rats performing a nonspatial radial maze task. J. Neurosci. *14*, 6553–6563.

Zeidman, P., and Maguire, E.A. (2016). Anterior hippocampus: the anatomy of perception, imagination and episodic memory. Nat. Rev. Neurosci. *17*, 173–182.

Zeithamova, D., Dominick, A.L., and Preston, A.R. (2012a). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. Neuron *75*, 168–179.

Zeithamova, D., Schlichting, M.L., and Preston, A.R. (2012b). The hippocampus and inferential reasoning: building memories to navigate future decisions. Front. Hum. Neurosci. *6*, 70–70.

Zhang, H., Zherdeva, K., and Ekstrom, A.D. (2014). Different "routes" to a cognitive map: dissociable forms of spatial knowledge derived from route and cartographic map learning. Mem. Cognit. *42*, 1106–1117.

SUMMARY

How can I remember going cross-country skiing for the first time? What mechanisms in my brain allow me to recall when and where I made my first steps on skis? Why can I, more than two years later, vividly picture myself on the slopes of the Bymarka near Trondheim, Norway, while sitting in my new apartment in Leipzig, Germany? In this thesis, I set out to investigate how our brain forms memories of events that occur at a specific place and time and how we can draw on these experiences to remember past events and imagine future scenarios. I explore the idea that key coding principles identified in spatial navigation research could operate across different cognitive domains, such as memory, to organize our experience. To address these questions, I focused on the hippocampus and the neighboring entorhinal cortex.

In the human brain, the hippocampus is situated deep in the medial temporal lobe. As described in **Chapter 1**, the hippocampus has been considered a crucial brain structure for so-called episodic memories ever since the famous patient H.M., whose hippocampus had been surgically removed. After this surgery, H.M. was unable to form new memories of events he encountered. Episodic memories are defined as vivid recollections of when in time and where in space events occurred. But how exactly do the hippocampus and the surrounding entorhinal cortex support our memory? Research on spatial navigation has shed light on fundamental coding principles of the hippocampal-entorhinal region and I summarize some key findings in **Chapter 1**. The ground-breaking discovery of place cells in the rat hippocampus led to the proposal that the hippocampus forms a cognitive map of our surroundings — a proposal that is highly influential till this day. A place cell is a neuron that fires preferentially when the animal is at one specific position in an environment. Different place cells fire at different positions in space, so that they collectively form a map-like representation of the environment. Extensive research has identified various other functionally defined cell types that contribute to the animal's cognitive map. Among them are head direction cells that convey compass-like information about the animal's head direction. Grid cells in the medial entorhinal cortex, one synapse upstream of the hippocampus, exhibit regular, symmetric firing patterns that are thought to provide a coordinate system or metric for the cognitive map. Intracranial recordings in patients undergoing presurgical screenings and studies on spatial navigation combining brain imaging and virtual reality have demonstrated that the human hippocampal-entorhinal region operates using comparable neural codes. Throughout my doctoral work presented in this thesis, I investigated how such coding principles might enable human memory and other cognitive functions.

Imagining a future scenario might be like simulating a trajectory using our cognitive map. In **Chapter 2**, I tested whether human imagination recruits spatial coding mechanisms. I trained participants to navigate a large-scale virtual reality city. Specifically, participants learned the names and locations of buildings and how to find their way between them. Then, while they underwent functional magnetic resonance imaging (fMRI), I asked participants to imagine what they would see when facing specific directions from positions in front of particular buildings. I analyzed the data using representational similarity analysis, which compares the similarity of activity patterns exhibited by local groups of voxels. In the parahippocampal gyrus, voxel patterns were more similar when participants imagined the same direction — an absolute directional code similar to the compass-like information conveyed by head direction cells. In the posterior-medial subregion of the entorhinal cortex, similarity patterns were characterized by a 60°-modulation. This resembles the 60°-symmetric firing patterns that are the hallmark of grid cells in the rodent medial entorhinal cortex, which is considered the homologue region of the human posterior-medial entorhinal subregion. Together, these findings suggest that picturing a scene in front of our mind's eye, for example when remembering a past event or imagining a future scenario, recruits some of the fundamental building blocks of our cognitive map.

Theoretical models of grid-cell function build on the regularity of grid firing patterns to encode positions and to compute distances and directions between them. However, studies have demonstrated that grid patterns can be distorted, for example through the geometry of the boundaries forming the environment that the animal navigates. In **Chapter 3**, I test whether spatial memories formed in an environment known to deform grid patterns in rodents are systematically distorted. In this behavioral experiment, participants navigated a square and a trapezoid environment using a highly immersive virtual reality system consisting of a head mounted display and a motion platform that translated their physical steps and rotations into virtual movement. In each environment, participants learned the positions of objects. Participants' memory for positions in the trapezoid was less precise than in the square and, within the trapezoid, memory was particularly degraded in its narrow end. This memory profile mirrors the severity of grid pattern distortions observed in rodents navigating a trapezoid. Further, I asked participants to estimate distances between learned positions outside of the trapezoid to test for persistent mnemonic distortions. If positions are encoded using a compressed or stretched grid pattern, then grid patterns change more or less as a function of the distance between two positions. Estimates of the distances between positions that were encoded using deformed

grid patterns should therefore be distorted outside of the trapezoid environment. Consistently, participants estimated identical distances to be different between the narrow and broad part of the trapezoid, an effect predicted based on the stronger distortions of grid patterns in the trapezoid's narrow end. In additional analyses, I reconstructed the individual positions participants remembered from their pairwise distance estimates to show that these reconstructed mnemonic maps explained object position memory in the trapezoid better than the true object positions. Collectively, the findings from this behavioral experiment suggest that human positional memory is subject to distortions through environmental geometry that can be predicted from the distortions of grid patterns observed in rodents; consistent with the idea that our memory for where events take place is supported by entorhinal grid coding.

Events take place not only at a certain position in space, but also at a specific moment in time. In **Chapter 4**, I investigated the question how the human entorhinal cortex could support our memory of the sequence in which events unfold over time. In this experiment, participants learned to navigate along a fixed route through a virtual city. Their task was to learn where and when during the traversal of the route events occurred. These events were defined by the encounter of objects along the path. Participants underwent fMRI before and after learning. During these scanning sessions, participants saw the same objects they encountered along the route, but in random order. This enabled me to quantify similarity changes of object representations from before to after learning. Object representations in the entorhinal cortex changed in a way that resembled the temporal structure of the object sequence: Objects that were encountered in temporal proximity got relatively more similar compared to objects far apart in time; resulting in a negative correlation of representational change and the temporal distance between objects. Importantly, this effect was specific to the anterior-lateral entorhinal subregion and specific to the temporal rather than spatial structure of the event sequence. Furthermore, I reconstructed the timeline of events from representational change. Additionally, participants in whom the temporal distance between objects correlated more strongly with representational change tended to successively reproduce objects learned to be nearby in time when trying to recall all objects in a post-scan memory test. This demonstrates that entorhinal representations of how a sequence of events unfolds relate to how we retrieve these events from memory.

Chapter 5 describes an overarching theoretical framework for how spatial coding principles in the hippocampal-entorhinal region could contribute to human cognitive function. In this work, I bridged findings from rodent electrophysiology and systems

neuroscience with human cognitive neuroscience and a theory from cognitive science. The ideas put forward in this chapter build on the theory of conceptual spaces developed by Peter Gärdenfors, where stimuli are arranged in spaces spanned by feature dimensions, which allows the description of concepts as convex region in these spaces. Based on recent evidence demonstrating that the hippocampal-entorhinal region processes experience using coding principles discovered in navigation research across various stimulus domains, I suggest that the hippocampal-entorhinal region might map cognitive spaces. Place cells might signal specific positions in a given cognitive space defined by stimulus feature dimensions, similarly to how they signal locations during navigation through physical space. Likewise, grid cells could provide a metric for cognitive spaces, encoding relational and distance information about stimulus positions in cognitive space. This framework incorporates other key findings regarding the coding principles employed by the hippocampal-entorhinal region to map space for navigation. For example, place cells in the hippocampus flexibly remap to form representations of different spaces, whereas entorhinal grid cells from the same module change their firing more coherently. Furthermore, place cells and grid cells encode spatial positions at different granularities as a function of their anatomical position, which might enable the representation of knowledge in cognitive spaces at different levels of detail. Additionally, sequences of place cell and grid cell firing can simulate trajectories through space, providing a way for knowledge stored in cognitive spaces to guide future behavior.

In **Chapter 6**, I summarize the findings of the experiments described in **Chapters 2-4** and discuss how they relate to the theoretical framework put forward in **Chapter 5**. The demonstration that absolute directional codes akin to head direction signals and grid-like entorhinal patterns can be observed during imagination, which I describe in **Chapter 2**, relates to the notion that activity sequences of spatially tuned cells underlie mental simulation, for example in service of navigational planning, but potentially also for the simulation of trajectories through abstract cognitive spaces. The deformations of human spatial memory through the geometry of environmental boundaries I report in **Chapter 3**, which putatively going back to distortions of entorhinal grid patterns, speak to the idea that the entorhinal grid system provides a coordinate system for space during navigation and potentially also for cognitive spaces spanned by different feature dimensions. Next to space, time constitutes a fundamental dimension along which our experience is organized. The representation of the temporal relationships of events in the entorhinal cortex that I describe in **Chapter 5** could underlie the way we remember the sequence in which our experience unfolded.

Collectively, the work I presented in this thesis investigated the coding mechanisms by which the hippocampal-entorhinal region maps our experience. Space and time can be seen as a scaffold for our experience or, in the context of cognitive spaces, as fundamental dimensions along which events are organized. The empirical work described in this thesis describes how coding principles discovered in research on spatial navigation can support our episodic memory. The hippocampal-entorhinal region might not only allow me to reminisce about past events such as my first adventure on cross-country skis, but, as proposed in the theoretical sections, provide a framework to organize experience for flexible cognition more broadly.

NEDERLANDSE SAMENVATTING

Hoe kan ik me herinneren aan mijn eerste keer langlaufen? Welke mechanismen in mijn hersenen stellen me in staat om me te herinneren waar en wanneer ik mijn eerste afdaling op de ski's heb gewaagd? Waarom kan ik me meer dan twee jaar later de hellingen van de Bymarka bij Trondheim (Noorwegen) levendig voor de geest halen, terwijl ik in mijn nieuwe appartement in Leipzig (Duitsland) zit? In dit proefschrift wil ik onderzoeken hoe onze hersenen herinneringen aan gebeurtenissen vormen, die zich op een bepaalde plaats en tijd voordoen, en hoe we ons op basis van deze ervaringen aan gebeurtenissen uit het verleden kunnen herinneren en ons toekomstscenario's kunnen voorstellen. Ik onderzoek het idee dat de belangrijkste coderingsprincipes die in ruimtelijke navigatieonderzoek zijn geïdentificeerd, in verschillende cognitieve domeinen – zoals het geheugen – actief zouden kunnen zijn om onze ervaringen te organiseren. Om deze vragen te kunnen beantwoorden heb ik me op de hippocampus en de aangrenzende entorinale cortex gericht.

De hippocampus bevindt zich diep in de mediale temporale kwab van het menselijk brein. Zoals beschreven in **hoofdstuk 1**, wordt de hippocampus beschouwd als een cruciale hersenstructuur voor de zogenaamde episodische herinneringen. De beroemde patiënt H.M., wiens hippocampus chirurgisch werd verwijderd, illustreerde de werking van de hippocampus. Na zijn operatie was H.M. niet meer in staat om nieuwe herinneringen te vormen aan gebeurtenissen in zijn leven. Episodische herinneringen worden gedefinieerd als levendige herinneringen met betrekking tot wanneer en in welke ruimte gebeurtenissen opgetreden zijn. Maar op welke manier ondersteunen de hippocampus en de aangrenzende entorinale cortex ons geheugen precies? Onderzoek naar ruimtelijke navigatie heeft duidelijkheid gebracht over de fundamentele coderingsprincipes van de hippocampus-entorinale regio. Enkele belangrijke bevindingen heb ik samengevat in **hoofdstuk 1**. De baanbrekende ontdekking van plaatscellen in de hippocampus inspireerde de theorie dat de hippocampus een cognitieve kaart van onze omgeving vormt – een voorstel dat tot op de dag van vandaag van grote invloed is. Een plaatscel is een neuron dat voornamelijk afvuurt wanneer een dier zich op één specifieke plaats binnen een omgeving bevindt. Verschillende plaatscellen geven signalen af op verschillende plaatsen binnen een ruimte om samen een kaartvormige weergave van de omgeving te genereren. Aan de hand van uitgebreid onderzoek zijn verschillende andere, functioneel gedefinieerde celtypes geïdentificeerd, die bijdragen tot de cognitieve kaart van een dier. Daaronder bevinden zich hoofdrichtcellen die kompasvormige informatie over de hoofdrichting van een dier overbrengen. Rastercellen in de

mediale entorinale cortex, één synaps stroomopwaarts van de hippocampus, vertonen regelmatige, symmetrische afvuurpatronen waarvan verondersteld wordt dat deze een coördinatenstelsel of metriek vormen voor de cognitieve kaart. Intracraniële opnames bij patiënten tijdens pre-operatieve screenings en studies over ruimtelijke navigatie in combinatie met hersenbeeldvorming en virtuele realiteit hebben aangetoond dat de menselijke hippocampus-entorinale regio werkt met behulp van vergelijkbare neurale codes. In mijn doctoraatsonderzoek dat in dit proefschrift wordt gepresenteerd heb ik onderzocht hoe dit soort coderingsprincipes het menselijk geheugen en andere cognitieve functies in staat zouden kunnen stellen om te functioneren.

Het voorstellen van een mogelijk toekomstscenario zou gelijkenissen kunnen vertonen aan het simuleren van een traject met behulp van onze cognitieve kaart. In **hoofdstuk 2** heb ik getest of de menselijke verbeelding gebruik maakt van ruimtelijke coderingsmechanismen. Ik heb de deelnemers getraind om te navigeren in een grootschalige, virtual reality-stad. De deelnemers leerden hierbij de namen en locaties van gebouwen en hoe ze daartussen hun weg konden vinden. Terwijl ze functionele magnetische resonantie beeldvorming (fMRI) ondergingen, vroeg ik de deelnemers om zich voor te stellen wat ze zouden zien als ze vanuit bepaalde posities voor bepaalde gebouwen en in specifieke richtingen zouden kijken. Ik analyseerde de gegevens met behulp van een representatieve overeenkomstenanalyse (RSA), die de gelijkenis van de activiteitspatronen van lokale voxelgroepen vergelijkt. In de gyrus parahippocampalis vertoonden de voxelpatronen meer gelijkenissen wanneer de deelnemers zich dezelfde richting voorstelden — een absolute richtingcode vergelijkbaar met de kompasvormige informatie die door de hoofdrichtingscellen wordt overgebracht. In de posterieure-mediale subregio van de entorinale cortex werden gelijkenispatronen gekenmerkt door een 60°-modulatie. Dit lijkt op de symmetrische afvuurpatronen van 60° die kenmerkend zijn voor rastercellen in de mediale entorinale cortex van knaagdieren. Dit wordt beschouwd als het homologe gebied van de menselijke posteriore-mediale entorinale subregio. Samen suggereren deze bevindingen dat bij onze hersenen bij het voorstelen van een bepaalde scène, zoals een herinnering van een gebeurtenis uit het verleden of het voorstellen van een toekomstscenario, enkele fundamentele bouwstenen van onze cognitieve kaart worden gebruikt.

Theoretische modellen van de rastercelfunctie bouwen voort op regelmatigheden in rasterafvuurpatronen om posities te coderen en afstanden en richtingen tussen deze posities te berekenen. Studies hebben echter aangetoond dat rasterpatronen vervormd

kunnen zijn, bijvoorbeeld door de geometrie van de grenzen van de omgeving waarin het dier navigeert. In **hoofdstuk 3** test ik of ruimtelijke geheugens die zijn gevormd in een omgeving waarvan bekend staan dat deze het rasterpatronen van knaagdieren vervormen ook systematisch vervormd zijn. In dit gedragsexperiment navigeerden deelnemers door een vierkant en een trapeziumvormige omgeving met behulp van een bijzonder meeslepend virtual-realitysysteem dat bestond uit een op het hoofd aangebrachte display en een bewegend platform dat hun fysieke stappen en rotaties vertaalde naar een virtuele beweging. In elke omgeving leerden de deelnemers de posities van objecten kennen. Het geheugen van de deelnemers voor posities in de trapezoïde was minder nauwkeurig dan voor het vierkant en binnen de trapezoïde was het geheugen in het smalle uiteinde bijzonder aangetast. Dit geheugenprofiel komt overeen met de mate van vervorming in de rasterpatronen die ook bij knaagdieren die door een trapezoïde navigeren waargenomen zijn. Verder vroeg ik de deelnemers om de afstanden tussen aangeleerde posities buiten de trapezoïde te schatten om te testen op blijvende mnemonische vervormingen. Als posities worden gecodeerd met een gecomprimeerd of uitgerekt rasterpatroon, dan veranderen de rasterpatronen min of meer als een functie van de afstand tussen twee posities. Schattingen van de afstanden tussen posities die gecodeerd zijn met behulp van vervormde rasterpatronen moeten daarom buiten de trapeziumvormige omgeving worden vervormd. De deelnemers schatten steeds identieke afstanden tussen het smalle en brede deel van de trapezoïde in, een effect dat voorspeld wordt op basis van de sterkere vervorming van de rasterpatronen in het smalle uiteinde van de trapezoïde. In aanvullende analyses heb ik de individuele posities die de deelnemers zich herinnerden van hun paarsgewijze afstandsschattingen gereconstrueerd om aan te tonen dat deze gereconstrueerde mnemonische kaarten het objectpositiegeheugen in de trapezoïde beter verklaarden dan de werkelijke objectposities. Gezamenlijk tonen de bevindingen van dit gedragsexperiment aan dat het menselijk positiegeheugen onderhevig is aan vervormingen door omgevingsgeometrie, die kunnen worden voorspeld op basis van de vervormingen van rasterpatronen die bij knaagdieren wordt waargenomen. Dit is in overeenstemming met het idee dat ons geheugen dat verantwoordelijk is voor de ruimtelijke duiding van gebeurtenissen wordt ondersteund door entorinale rastercodering.

Gebeurtenissen doen zich niet alleen voor op een bepaalde positie in de ruimte, maar ook op een bepaald moment in de tijd. In **hoofdstuk 4** onderzocht ik de vraag hoe de menselijke entorinale cortex onze herinnering van de volgorde waarin gebeurtenissen zich afspelen kan ondersteunen. In dit experiment leerden de deelnemers te navigeren langs een vastgelegde route door een virtuele stad. Hun taak

was om te leren waar en wanneer er langs de route gebeurtenissen plaatsvonden. Bepaalde momenten waarin deelnemers objecten langs de route tegenkwamen werden als gebeurtenissen vastgelegd. De deelnemers ondergingen een fMRI voor en na het leren. Tijdens deze scansessies zagen de deelnemers dezelfde objecten die ze langs de route tegenkwamen, maar in willekeurige volgorde. Dit stelde me in staat om veranderingen in overeenkomsten tussen de objectvoorstellingen van voor tot na het leren te kwantificeren. Objectvoorstellingen in de entorinale cortex veranderden op een manier die gelijkenissen vertoonde aan de temporele structuur van de objectsequentie: objecten die deelnemers kort op elkaar aantroffen vertoonden onderling grotere gelijkenissen dan objecten waartussen meer tijd lag. Dit resulteerde in een negatieve correlatie tussen representatieveranderingen en de tijdsverschillen tussen objecten. Hierbij is het belangrijk om te noemen dat dit effect specifiek van toepassing was op de anterieure-laterale entorinale subregio en specifiek voor de temporele in plaats van de ruimtelijke structuur van de gebeurtenissequentie. Daarnaast heb ik de tijdlijn van de gebeurtenissen met behulp van de representatieve verandering. Deelnemers bij wie de tijdsafstand tussen objecten sterker correleerden met representatieveranderingen hadden daarnaast de neiging om objecten waarvan ze geleerd hadden dat deze temporaal gezien dicht bij elkaar lagen successief te reproduceren wanneer ze alle objecten in een na-scan geheugentest probeerden op te roepen. Dit toont aan dat entorinale representaties van het afspelen van een sequentie van gebeurtenissen verbonden is aan met hoe we deze gebeurtenissen oproepen uit ons geheugen.

Hoofdstuk 5 beschrijft een overkoepelend, theoretisch kader voor de wijze waarop ruimtelijke coderingsprincipes in het hippocampus-entorinale gebied kunnen bijdragen aan de menselijke cognitieve functie. In dit werk overbrugde ik de bevindingen uit de knaagdierelektrofysiologie en systeemneurologie met de menselijke cognitieve neurowetenschap en een theorie uit de cognitieve wetenschap. De ideeën in dit hoofdstuk bouwen voort op de theorie van conceptuele ruimtes ontwikkeld door Peter Gärdenfors. Hierin zijn gerangschikt in ruimtes die verbonden worden door functiedimensies. Dit maakt het mogelijk om concepten als convexe regio's in deze ruimtes te beschrijven. Op basis van recent bewijs dat aantoont dat de hippocampus-entorinale regio ervaringen verwerkt in verschillende stimulusdomeinen met behulp van coderingsprincipes uit navigatieonderzoek, opper ik de theorie dat de hippocampus-entorinale regio cognitieve ruimten in kaart kan brengen. Plaatscellen kunnen mogelijk specifieke posities aanduiden in een bepaalde cognitieve ruimte gedefinieerd door stimulusfunctiedimensies, op dezelfde manier waarop ze locaties doorgeven tijdens het navigeren door een fysieke

ruimte. Evenzo kunnen rastercellen een metriek vormen voor cognitieve ruimtes, die relationele en afstandsinformatie over stimulusposities in de cognitieve ruimte coderen. Dit kader omvat andere belangrijke bevindingen met betrekking tot de coderingsprincipes die door de hippocampus-entorinale regio worden toegepast om navigatieruimten in kaart te brengen. Plaatscellen in de hippocampus kunnen zichzelf bijvoorbeeld flexibel binnen een kader omvormen om zo representaties van verschillende ruimtes te vormen, terwijl entorinale rastercellen uit dezelfde module zichzelf aan kunnen passen om meer coherent af te vuren. Bovendien coderen plaatscellen en rastercellen ruimtelijke posities met verschillende granulariteiten als een functie van hun anatomische positie, wat de weergave van kennis in cognitieve ruimtes op verschillende detailniveaus mogelijk zou kunnen maken. Bovendien kunnen afvuursequenties van plaats- en rastercellen ruimtelijke trajecten simuleren. Op deze manier kan kennis die is opgeslagen in cognitieve ruimtes toekomstig gedrag sturen.

In **hoofdstuk 6** vat ik de bevindingen van de in **hoofdstuk 2-4** beschreven experimenten samen en bespreek ik hoe deze zich verhouden tot het in **hoofdstuk 5** voorgestelde theoretische kader. In **hoofdstuk 2** wordt beschreven hoe ik aantoonde dat absolute richtingscodes die vergelijkbaar zijn met hoofdrichtingssignalen en rastervormige, entorinale patronen waargenomen kunnen worden tijdens mentale verbeeldingen. Dit heeft te maken met het onderliggende concept dat activiteitenreeksen van ruimtelijk afgestemde cellen ten grondslag liggen aan mentale simulatie, bijvoorbeeld voor navigatieplanning, maar mogelijk ook voor de simulatie van trajecten door abstracte cognitieve ruimtes. De vervormingen van het menselijk ruimtelijk geheugen door de geometrie van omgevingsgrenzen die ik in **hoofdstuk 3** presenteer en die op een bepaalde manier teruggrijpen naar vervormingen van entorinale rasterpatronen, ondersteunen het idee dat het entorinale rastersysteem een ruimtelijk coördinatensysteem biedt tijdens navigeren en mogelijk daarnaast ook voor cognitieve ruimtes die door verschillende functiedimensies worden overkoepeld. Naast ruimte vormt tijd een fundamentele dimensie om onze ervaring te organiseren. De weergave van de temporele relaties van gebeurtenissen in de entorinale cortex die ik in **hoofdstuk 5** beschrijf, zou de basis kunnen vormen voor de manier waarop we ons de volgorde herinneren waarin onze ervaring zich heeft voorgedaan.

Het samenvattend werk dat ik in dit proefschrift heb gepresenteerd onderzocht de coderingsmechanismen waarmee de hippocampus-entorinale regio onze ervaring in kaart brengt. Ruimte en tijd kunnen worden gezien als een raamwerk voor onze

ervaring of, in de context van cognitieve ruimtes, als fundamentele dimensies voor het organiseren van gebeurtenissen. Het empirische werk in dit proefschrift beschrijft hoe coderingsprincipes, die werden ontdekt in het onderzoek naar ruimtelijke navigatie, ons episodisch geheugen kunnen ondersteunen. Met behulp van de hippocampus-entorinale regio kan ik niet alleen herinneringen ophalen aan gebeurtenissen uit het verleden, zoals mijn eerste avontuur op langlaufski's, maar biedt, zoals voorgesteld in de theoretische gedeelten, een kader om de ervaring van flexibele cognitie te organiseren in een bredere context.

ACKNOWLEDGEMENTS

Dear reader, thank you for your interest in my thesis. The work presented in this thesis would not have been possible without the support of many amazing people. There are too many to name all of you, but I am grateful to each and every one of you. It is not easy to find the right words to say special thanks to few of you.

Christian, I would like to thank you for your continued support throughout my scientific journey that involved institutes in three countries by now. Your optimism and enthusiasm are contagious and have helped me overcome the challenges of pursuing a PhD. I am grateful for the many opportunities you have created for me over the years and the thoughtful advice you have given me. I am looking forward to continue to work with you.

Lorena, thank you for taking me by the hand during the first months and years at the Donders and making the start into the PhD adventure so awesome. I benefited from your supervision tremendously and learned so much from you. Thanks for all the advice you have given me as a scientist and friend! I am excited that we still find (virtual) time to continue our collaborations.

Caswell, when I met you at the Cumberland Lodge, you immediately inspired me as a scientist and incredibly bright mind. It has been a great pleasure to work with you on the trapezoid project and I look forward to future collaborations. **Will**, your modeling work provided a crucial piece of the puzzle for the trapezoid project (even if it is not presented in this thesis). Thanks for that! **Peter** and **Edvard**, many thanks for your contributions to our review paper. It has been a pleasure to work with you on this manuscript and I am grateful to have had the opportunity to benefit from your insight and experience. Hopefully, we can continue our fruitful collaborations in the future.

Thinking back to the start of my time at the Donders: **Monja**, we became friends soon after we first became office mates in our tiny office next to the kitchen. As our friendship grew, we also moved to our new, larger office. I always love(d) sharing my thoughts on everything science and non-science with you. Thanks for being such a good friend and for being my paranymph for the defense of this thesis!

Nils and **Catalina**, thanks for being such great office mates; I still think back to all the fun times we had in the office and in Nijmegen. Thanks for sharing all the ups

and downs of (PhD) life with me. I loved participating in bake-offs with you and still remember the cloak-and-dagger operation bringing the big red couch into our possession, without which our office would not have been the same. Thanks also to my office mates in Trondheim and Leipzig: **Mona**, **Raphael** and **Misun**, it has been (and continues to be) awesome to share a room with you!

I would like to thank the members of the Doellerlab for making it such a stimulating group to work in. **Silvy**, thanks for being a great companion and friend. Sharing the excitement and the occasional frustration of our PhD endeavor over our morning coffees was wonderful. **Alex** and **Sander**, you showed me it was okay to obsess over figure and poster designs and the importance of a good six-letter acronym (I am taking continuous efforts to keep this tradition alive). **Lonja**, you're a great friend and can't thank you enough for all the weekends you spent with me at the 7T in Essen (Mr. Wasabi says hi). **Naomi**, thanks for always sharing your critical views and the passionate discussions we had. **Branka**, **Stephanie**, **Loes**, **Staudi** (finally some more football expertise in the lab, but then a Bayern fan…), **Thomas**, **Sasha**, **Isabella**, and **David**, thanks for being such great colleagues at the Donders and making the lab such an inspiring place to work in. **Meryl**, thanks for making the lab a funnier place with your great eye for the weird and strange side of things (together with **Lonja** and **Lorena** of course).

The Donders is a very special place, especially because of all the amazing people who make sure everything is running so smoothly. Thank you very much for making this a great institute. This includes the **administration**, thanks especially to **Ayse**, **Sandra** and **Nicole** for making things work even during my years as an external PhD candidate after leaving Nijmegen. **Tildie**, you are one of the keys to this institute's success. Thank you for always being helpful! I further owe the **technical group** at the Donders many thanks: You are amazing and have established a culture of well-thought-out procedures to do great science, from data collection (special thanks to **Paul**, guardian of the MR scanners!), high-performance computing to data archiving and sharing. Leaving the Donders made me realize (even more) just how well things work at this institute because of you. Thank you! The Donders is a particularly great place because of the **Social Donderians** (there are too many to name, you know who you are!). I will always remember it as an extremely helpful and fun environment. Thanks also to the **Donders football team** (**Eelco**, **Sean**, **Ruud**, **Alan**, **Tom**, **Pim**, **Maarten**, **René**, **Izabela**, **Monja** and many more).

Kevin, thanks for being such a good housemate during my time in Nijmegen. I have fond memories of the time in our apartment and all the delicious meals we shared! This also includes **Christian** of course!

A very special thanks goes to **Tobias** and **Matthias**. **Tobias**, you were the first to show me around the Donders and you being my paranymph for the defense of this thesis closes this circle. Thanks for all your advice over the years. Your balanced views and calm presence have always helped me focus on what's important. **Matthias**, you are one of the most dedicated scientists I know. Many of my projects improved through our discussions and I love sharing thoughts on science and life with you. Both of you, I will never forget moving to Norway with you and being the first generation at the Kavli. This was such an exciting time. Having you as friends and colleagues made it incredibly amazing!

Soon the lab in Trondheim grew and we were joined by many amazing people: **Dörte**, **Markus**, **Nacho**, **Anne**, **Lilith**, **Annelene**, **Josh**, **Merethe**, **Gøril**, **Ingrid**, **Sabine** and **Raphael**, thanks for being a wonderful team of friends and colleagues in Trondheim! I loved my time in Norway with you. Thanks also to everyone at the **Kavli**. There are so many fantastic scientists (way too many to name all of you) at this institute and I have learned so much about many of the different facets of neuroscience from you.

Over the years, I have had the great pleasure to work with several brilliant students. **Jackeline**, **Tom**, **Anne**, **Britt** and **Natalie**, it has been a privilege to work on projects with you and see you grow as scientists. **Dörte**, the same goes for you of course and I am excited to work with you. I am looking forward to see where our projects will take us in the future.

Mona and **Misun**, you made moving to the MPI easy for me. I admire both of you as bright and dedicated scientists. I am very lucky to have you as friends and colleagues in Leipzig and I am beyond excited to have started the next adventure with you.

We were soon joined by new members of the lab at the MPI: **Volker**, **Stephanie**, **Felix**, **Theo**, **Blanca**, **Artem**, **Marco**, **Alex** and **Johanna**, I am much looking forward to (keep) working with you. **Susanne**, **Kerstin** and **Anja**, I would like to thank you for all the support you have provided to make settling in at the new institute and the new city a pleasure.

Die wichtigsten Personen kommen natürlich zum Schluss. Meine liebe **Anja**, vielen Dank für Deine fortwährende Unterstützung in den letzten Jahren. Dass Du mich auf dieser Reise in Richtung Promotion durch drei Länder begleitet hast, ist wunderbar und unglaublich mutig von Dir. Ich bin dankbar, dass Du Verständnis für die Anforderungen der Wissenschaft hast und ich immer auf deinen Rückhalt zählen kann. Ich kann nicht sagen, wie glücklich ich bin Dich in meinem Leben zu haben!

Ich danke meinen **Freundinnen und Freunden**, die während der Erstellung dieser Arbeit immer für mich da waren, für ihre Unterstützung. Danke Euch allen! This includes my English-speaking friends as well of course. Many thanks for your support!

Auch meinen Eltern und meinem Bruder möchte ich danken für ihre Unterstützung und Ermutigungen auf diesem Weg. Liebe **Mama**, lieber **Papa**, lieber **Mauro**, ich bin sehr dankbar, dass Ihr mich immer ermuntert habt Neues auszuprobieren und meinen Weg zu finden. Wohin mein Weg auch führt, so weiß ich, dass ich bei Euch immer eine sichere Basis habe. Liebe **Oma**, auch Dir möchte ich danken. Deine Frage, wie wir das Gedächtnis besser machen können, habe ich mit dieser Arbeit noch nicht beantwortet, aber vielleicht kommt das ja als Nächstes!

CURRICULUM VITAE

Jacob Lukas Sarid Bellmund was born on 1 January 1988 in Adelebsen, Germany. After graduating from Albion High School (Michigan, USA) and the Theodor-Heuss-Gynmasium in Göttingen, Germany, Jacob pursued the study of Psychology at the Philipps-University Marburg, Germany. During his studies, he worked as a student assistant with Prof. dr. Patrick Khader and Dr. Mira-Lynn Chavanon. Driven by a strong interest in the cognitive neuroscience of episodic memory, Jacob completed a research internship in the laboratory of Prof. dr. Charan Ranganath at the University of California, Davis (USA). Jacob obtained his Diplom (M.Sc. equivalent) in Psychology in 2013.

In November 2013, Jacob started as a PhD student under the supervision of Prof. dr. Christian Doeller and Dr. Lorena Deuker at the Donders Institute for Brain, Cognition, and Behaviour. Over the following years, Jacob investigated how the hippocampal-entorhinal region supports our memory and imagination. From September 2016, he continued to work towards his PhD at the Kavli Institute for Systems Neuroscience | Centre for Neural Computation at the Norwegian University of Science and Technology in Trondheim, Norway. Since December 2018, Jacob works at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany.

Jacob is passionate about data visualization and using graphics to communicate scientific ideas. He seeks answers to the question how scientific practice can be improved. Further, Jacob is interested in science communication and is the curator of the lab's twitter account with currently around 3700 followers. His research has been covered by major news outlets such as Forbes Magazine, ScienceDaily, Wissenschaft.de and The Economist's Espresso App as well as long-form science magazines such as Quanta Magazine and Nautilus Magazine. Recently, he has written an article to be published in the German magazine Gehirn & Geist to share insights from his research with a broad audience.

LIST OF PUBLICATIONS

Peer-reviewed publications

J. L. S. Bellmund, L. Deuker, T. Navarro Schröder, C. F. Doeller (2016). Grid-cell representations in mental simulation. *eLife*, *5*, e17089. https://doi.org/10.7554/eLife.17089

L. Deuker, **J. L. S. Bellmund**, T. Navarro Schröder, C. F. Doeller (2016). An event map of memory space in the hippocampus. *eLife*, *5*, e16534. https://doi.org/10.7554/eLife.16534

M. Nau, T. Navarro Schröder, **J. L. S. Bellmund**, C. F. Doeller (2018). Hexadirectional coding of visual space in human entorhinal cortex. *Nature Neuroscience*, *21* (2), 188–190. https://doi.org/10.1038/s41593-017-0050-8

J. L. S. Bellmund, P. Gärdenfors, E. I. Moser, C. F. Doeller (2018). Navigating cognition: Spatial codes for human thinking. *Science*, *362*, eaat6766. https://doi.org/10.1126/science.aat6766

J. L. S. Bellmund, L. Deuker, C. F. Doeller (in press). Mapping sequence structure in the human lateral entorhinal cortex. *eLife*. Preprint available at https://doi.org/10.1101/458133

J. L. S. Bellmund, W. de Cothi, T. A. Ruiter, M. Nau, C. Barry, C. F. Doeller (accepted for publication). Deforming the metric of cognitive maps distorts memory. *Nature Human Behaviour*. Preprint available at https://doi.org/10.1101/391201

Resources

J. L. S. Bellmund, L. Deuker, C. F. Doeller (2018). Donderstown. *Open Science Framework*. https://doi.org/10.17605/osf.io/78uph

DONDERS GRADUATE SCHOOL FOR COGNITIVE NEUROSCIENCE

For a successful research Institute, it is vital to train the next generation of young scientists. To achieve this goal, the Donders Institute for Brain, Cognition and Behaviour established the Donders Graduate School for Cognitive Neuroscience (DGCN), which was officially recognised as a national graduate school in 2009. The Graduate School covers training at both Master's and PhD level and provides an excellent educational context fully aligned with the research programme of the Donders Institute.

The school successfully attracts highly talented national and international students in biology, physics, psycholinguistics, psychology, behavioral science, medicine and related disciplines. Selective admission and assessment centers guarantee the enrolment of the best and most motivated students.

The DGCN tracks the career of PhD graduates carefully. More than 50% of PhD alumni show a continuation in academia with postdoc positions at top institutes worldwide, e.g. Stanford University, University of Oxford, University of Cambridge, UCL London, MPI Leipzig, Hanyang University in South Korea, NTNU Norway, University of Illinois, North Western University, Northeastern University in Boston, ETH Zürich, University of Vienna etc. Positions outside academia spread among the following sectors: specialists in a medical environment, mainly in genetics, geriatrics, psychiatry and neurology. Specialists in a psychological environment, e.g. as specialist in neuropsychology, psychological diagnostics or therapy. Positions in higher education as coordinators or lecturers. A smaller percentage enters business as research consultants, analysts or head of research and development. Fewer graduates stay in a research environment as lab coordinators, technical support or policy advisors. Upcoming possibilities are positions in the IT sector and management position in pharmaceutical industry. In general, the PhDs graduates almost invariably continue with high-quality positions that play an important role in our knowledge economy.

For more information on the DGCN as well as past and upcoming defenses please visit: http://www.ru.nl/donders/graduate-school/phd/

DONDERS INSTITUTE

Radboud University Radboudumc